DOCUMENT RESUME

ED 027 069 PS 001 540

By-Walls, Richard T.
Cognitive Factors in Semantic Conditioning. A Thesis in Educational Psychology.
Pennsylvania State Univ., University Park. Dept. of Educational Psychology.
Spons Agency-National Inst. of Child Health and Human Development, Bethesda, Md.
Pub Date Dec 68
Note-143p.
EDRS Price MF $0.75 HC-$7.25
Descriptors-*Behavioral Science Research, *Children, Cognitive Development, *Cognitive Processes, *Conditioned Response, Discrimination Learning, Grade 1, Operant Conditioning, Orthographic Symbols, Perception, Pictorial Stimuli, Reinforcement, Semantics, Stimulus Behavior, *Verbal Operant Conditioning

One control group and eight experimental groups, each composed of 12 first grade children, participated in this experiment. It was designed to investigate the effects of frequency of reinforcement and repeated evaluation of stimuli on the conditioning of preferences. Each child participated in the experiment for seven consecutive school days. The experimental children were divided into two stimulus condition groups; one received the same stimuli each session, and one received different stimuli from one session to the next. Each of these two groups was also divided into four rating groups. During a rating session, the child was asked to choose from 14 adjectives the one that he thought best described a particular Greek letter. A conditioning session consisted of a task involving reinforcement of particular Greek letters. The overall procedure allowed for the determination of any subject's change in preference for certain Greek letters as a function of either (1) number of evaluation sessions or (2) frequency of reinforcement. It was found that both factors affected preference. (WD)

ED0 27069

U. S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

The Pennsylvania State University

The Graduate School

Department of Educational Psychology

Cognitive Factors in Semantic Conditioning

A Thesis in

Educational Psychology

by

Richard T. Walls

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

December, 1968

Date of Approval:

______________ ______________
Professor of Education and Psychology
Thesis Adviser

______________ ______________
Head of the Department of Educational
Psychology

PS 001540

ACKNOWLEDGMENTS

The research reported herein was supported by Grant Number HD-02025 from the National Institute of Child Health and Human Development, U. S. Public Health Service, Francis J. Di Vesta, Principal Investigator.

Primary consultation for this research was provided at all phases by Professor Francis J. Di Vesta. The writer is deeply indebted for his efforts in painstaking guidance and criticism in the planning and execution of this project, and for his help and encouragement during the previous four years.

In addition, appreciation is given to the members of the thesis committee, Dr. Leon Gorlow, Dr. John F. Hall, Dr. Robert L. Lathrop, and Dr. William Rabinowitz, for their helpful advice and thoughtful reading at various stages of the work.

Thanks are also due to Andrew J. Chatlas (Director of Elementary Education in the Bellefonte Area School District), Ott B. Kantner (Principal of the Bellefonte Elementary School), and to the first-grade teachers, Eleanor Bathgate, Irmgard Cistone, Margret Long, and Gail Pohl for their cordial cooperation and support in the conduct of the present study. Appreciation is also expressed to the staff of the Pleasant Gap Elementary School for their cooperation during the pilot stages of the investigation.

TABLE OF CONTENTS

Page

Acknowledgments ii

List of Tables v

List of Figures vii

I. INTRODUCTION 1

II. REVIEW OF LITERATURE 10
- Automatic Strengthening and Cognitive Influences on Conditioning Involving Verbal Materials 10
  - The Law of Effect 10
  - Recent Investigations of Learning Without Awareness 13
  - Awareness as it Relates to Verbal Conditioning 17
  - Criteria for Awareness 20
  - Semantic Conditioning of Meaning 25
- The Influence of Mediated Experimental Demand Perceptions on Theory and Experimentation . 28
  - The Relation of Mediation to Awareness . . 28
  - The Perception of Experimental Demands . . 37
  - A Theoretical Structure for Assessing Perceived Demands 39
  - The Relevance of a Recent Series of Conditioning of Meaning Experiments . 43
- Summary 45

III. PURPOSE OF THE EXPERIMENT 48
  - Hypotheses 51
- Summary 57

IV. METHOD 59
  - Subjects 60
  - Apparatus, Rating Materials, and Filler Materials 62
    - Conditioning apparatus 62
    - Rating materials 65
    - Filler materials 65
  - Stimulus Materials 67
  - Measures 68
    - Adjective rating 68
    - Verbal forced-choice rating 69
    - Postexperimental awareness measures . 69
  - Procedure 70

Page

V. RESULTS 76
Postexperimental Awareness Data 88

VI. DISCUSSION 94
Reinforcement Procedures in the Development of Preferences 94
Rating and Preferences 97
Awareness and Preferences 101

VII. SUMMARY 104

REFERENCES 112

APPENDIX 122

LIST OF TABLES

Page

I. Some Possible Effects in a Long-Term Semantic Conditioning Experiment 52

II. Representation of Treatment Conditions 61

III. Four Sets of Stimuli for Conditioning and Rating . 66

IV. Analysis of Variance of Final Preference Scores . . 77

V. Means and Standard Deviations of Final Preference Scores 78

VI. Multiple Comparisons of Final Preference Scores by All Groups 81

VII. Analysis of Variance of Final Preference Scores for Stimulus and Rating Schedule Conditions 82

VIII. Mixed Analysis of Variance of Preference Scores by Groups SS-R4, DS-R4, and SS-EO Across Four Rating Trials 85

IX. Mean Preference Scores for All Groups on All Rating Days 86

X. Comparison of Final Mean Preference Scores for the Extra-Experimental Stimulus and the Standard Stimulus 89

XI. Weighted Means for Postexperimental Awareness Measures (Scores of 0, 1, 2) Compared with Means of Final Preference Score 91

XII. Number of Pupils per School Section (1, 2, 3, 4) of First Grade Children Among Experimental Conditions 126

XIII. Mean Age in Months of _S_s in Experimental Conditions 127

XIV. Adjectives and Bi-polar Scales Used in Rating Procedure 128

Page

XV. Mean Five Step Semantic Rating of Concepts in Pilot Evaluation 129

XVI. Mean Days Missed (Including Weekend) Between Experimental Sessions 130

LIST OF FIGURES

Page

1. Effects of conditioning sessions on verbal evaluation of the rewarded stimulus (Nunnally, Stevens, & Hall, 1965) 6

2. (A) Development of a sign; (B) Development of an assign (Osgood, Suci, & Tannenbaum, 1957) 30

3. Simplified descriptive schema for the development of mediated generalization (Cofer & Foley, 1942) . . . 32

4. The surface of the conditioning table 63

5. (a) Mean preference scores for SS groups and for the Nunnally, Stevens, and Hall (1965) group over all rating days. The latter group is comparable to the SS-R4 group in the present study 79

   (b) Mean preference scores for the DS groups and for the Control (SS-R0) group 79

6. Letter of information and thanks to parents, given to S on day 7 123

7. Sample computation of Jonckheere (1954) S statistic and t conversion 124

CHAPTER I

INTRODUCTION

In the past fifteen years considerable impetus has developed for the study of the role of verbal materials in conditioning and the influence of cognitive processes upon such conditioning. As a consequence, there now exists a large literature related to verbal operant conditioning. The earlier studies dealt with questions about types of responses and reinforcers (e.g., plural nouns or opinion statements rewarded by "mmm-hmm," "good," or a light); the later work concentrated on the question of whether verbal conditioning can occur without _S_s' awareness of the response-reinforcement contingency. For classical or semantic conditioning, with which this paper is primarily concerned, the body of literature is much more limited.

The role of _S_'s awareness is as important in semantic conditioning experiments as it is in conditioning of other responses, although it is true that relatively few studies have used verbal materials for both the CS and UCS in the classical conditioning paradigm. One class of such experiments involves the classical conditioning of meaning (Staats & Staats, 1959). The procedural paradigm consists of pairing a nonsense syllable (CS) with meaningful words (UCS) of positive or negative evaluation on the premise that meaning responses elicited by such words can be conditioned to a contiguously presented neutral stimulus. Again, early studies (Staats & Staats, 1957; Staats, Staats, & Biggs, 1958) reported no evidence of awareness and in general, deemphasized the role of cognitive processes while some subsequent investigations (Cohen, 1964) found no evidence for learning without awareness.

Another kind of semantic conditioning experiment involves the pairing of neutral objects with rewards (Nunnally, Duchnowski, & Parker, 1965). This body of research appears to be related to secondary reinforcement; although, it differs from the usual study of secondary reinforcement with respect to the dependent measures used. Traditionally, research involving the association of reward with neutral stimuli has been limited to investigating influences of a formerly neutral stimulus on learning and extinction of instrumental motor responses. Or, it has focused on the learning of new discriminations in which the stimulus serves as a cue related to obtaining a reward or to avoiding punishment (Nunnally & Faw, 1968). However, recent research suggests that the affective value of the formerly neutral stimulus is altered through association with reward.

When a S is consistently rewarded each time a particular neutral stimulus occurs, there are reliable effects on verbal evaluation of the stimulus, perceptual responses, and instrumental responses. For example, in one experiment by Nunnally, Stevens, and Hall (1965) a child sets in motion an apparatus which automatically selects a stimulus from a number of neutral geometric figures. Thus, the S initiates a selective action but is not the causal agent in that selection. The fortuitous occurrence of a randomly selected "pay-off" stimulus results in a reward such as candy, pennies, or marbles. If the reward stimulus is valued by the S, then, a degree of satisfyingness should also become associated with a contiguously presented neutral stimulus; that is, theories of secondary reinforcement suggest that a rewarded stimulus should come to elicit some of the same responses as are made to the reward. Thus, if reinforcement is preferred to nonreinforcement by S, a reinforced stimulus should come to be preferred to nonreinforced stimuli.

Although awareness has not been closely examined in the conditioning of meaning paradigm, there are several possible levels of awareness: (a) A S, in isolated cases, may not be aware that a reward is being given. A reinforcing stimulus may occur without S connecting it with any part of the learning task. Indeed, in some experiments he even may not notice them (e.g., the tap of a pencil or the flash of a small light). (b) A person can be aware of the regular occurrence of reinforcement without understanding the response contigency with which it is associated. Similarly, a child may know that an M and M candy is sometimes given to him but may never recognize other stimuli with which it is regularly associated; that is, he is unaware of the antecedent stimulus or stimuli. (c) Finally, the S may be aware of the reinforcing stimulus, associated stimuli, and/or the response contingency; that is, S knows that the reinforcing stimulus appears on each occasion that a given stimulus appears or that a given response is made. Thus, lack of awareness by S may be with respect to "the behavior itself . . . the relation of his behavior to some contingent event . . . the sensory experiences that usually accompany a given kind of stimulation . . . the fact that he is responding differentially to different stimuli . . . or contingencies in the environment which produce changes in S's behavior" (Adams, 1957, pp. 383-384).

In behavioristic terms, such concepts as awareness and S's perception of experimental demands are defined constructs. The antecedent conditions are the stimulus variations, including instructions given the S for conditioning, any events related to evaluations that might be required of the S, and the practice he has in perceiving, rating, or otherwise experiencing the stimuli. The consequent conditions are verbal reports or other behavioral results of manipulations deliberately designed for the purpose of identifying (observing) such relations. However, some investigators may have confounded such influences with the influence of practice, warm-up, or learning-to-learn on the final evaluation of whether conditioning actually occurred or was merely the result of Ss' awareness and learning to conform to E's demands.

A case in point is the Nunnally, Stevens, and Hall (Exp. I, 1965) procedure which is also the basis for the present investigation. They hypothesized that a given stimulus, when associated with a reward, would gain in preference value. Each S participated in the study for seven sessions over as many consecutive school days. Verbal evaluations (ratings) of the geometric stimuli were taken on the first, third, fifth, and seventh sessions. Conditioning trials (reinforcements for a "pay-off" stimulus) were administered on the second, fourth, and sixth sessions. The operations performed by the apparatus resembled those of a slot machine. Thus, when the S pulled a knob, a motor turned a hexagonal wheel so that six neutral geometric figures appeared one after another in a small aperture. One of these figures was assigned as a pay-off stimulus for a given child. Then, in the course of the experiment, he received a penny each time the wheel came to rest with that form showing in the window. In any one session (day 2, 4, or 6) S was allowed to play the slot machine for 25 minutes and received an average of 16 pennies for each conditioning session. The general finding was that preference (as measured by the number of positive and negative adjectives assigned by S) for the rewarded CS increases linearly over conditioning trials (Figure 1). However, it is important to note that the increase in preference might have been due to the demand factors (e.g., S hypothesizing that "E wants the rewarded one to be called good") associated with repeated ratings of the stimuli or, conditioning per se. With repeated ratings, as opposed to a single evaluation, there is a greater opportunity to form hypotheses about experimental demands. Since there was no control for the possible influence of repeated ratings it is not possible to separate these two effects.

Sudden insight into the nature of the response-reinforcement contingency is not necessarily manifested in a sharp increase in indices of learning, whether such indices consist of number of correct responses or of increase in preference for the pay-off stimulus. Nor is such awareness easily detectable from traditional dependent measures, designs, or postexperimental interviews. To reach high performance levels, practice at applying the principle

PS 001540

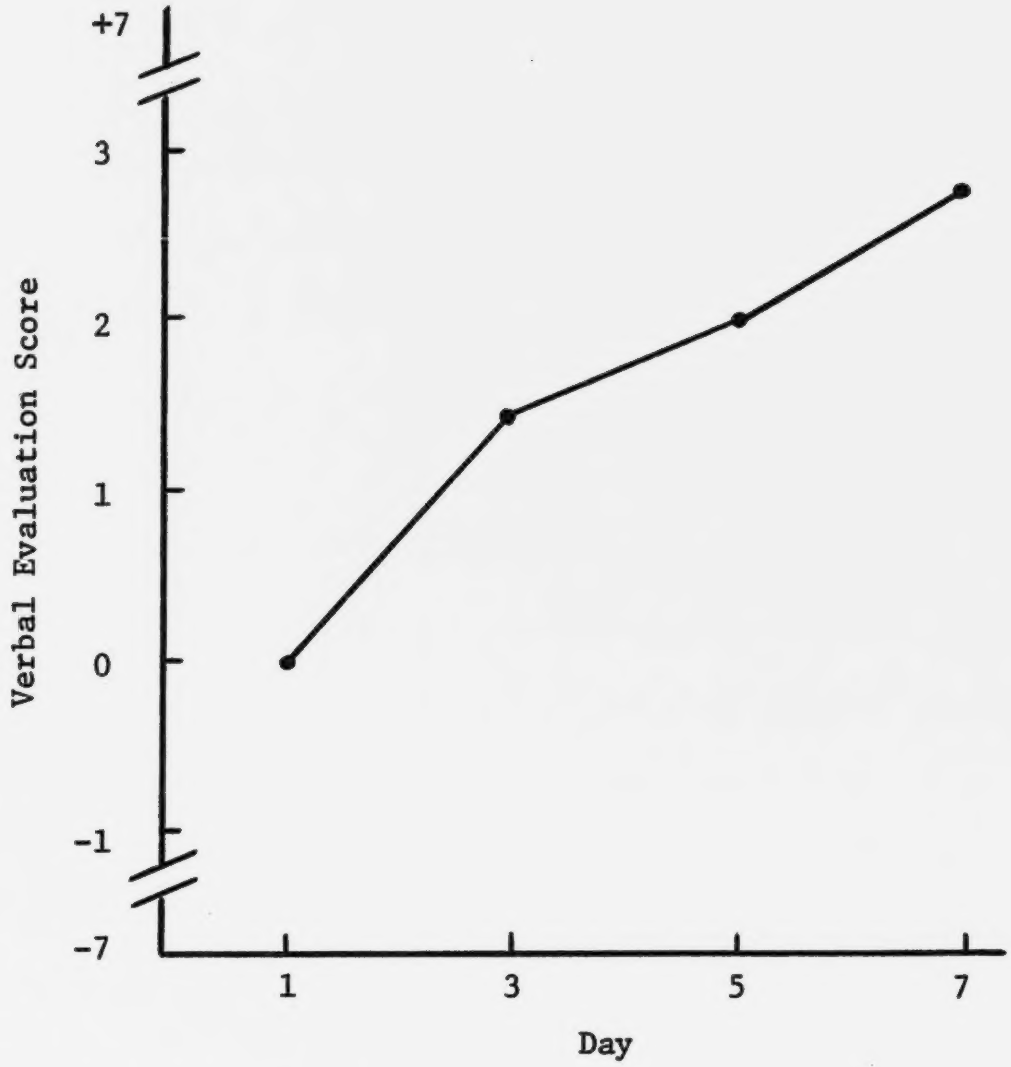


Figure 1. Effects of conditioning sessions on verbal evaluation of the rewarded stimulus (Nunnally, Stevens, & Hall, 1965).

may be as important as understanding the nature of the reinforcement, as learning the response, and/or as identifying the contingency principle (Postman & Jarrett, 1952).

When conducting research on verbal conditioning and learning without awareness in which very young children are employed, the difficulty in identifying awareness is considerably increased compared to situations in which adults are used. In the first place, children's language facility for these purposes may not be as efficient as those of adults. Secondly, it is difficult to know precisely what a particular set of experimental conditions conveys to a young child. The important question here is whether, by specifying what we mean by awareness and by defining the contents of the response and the procedures used to evoke it, one can identify the variables underlying verbal conditioning of meaning in such experiments as that conducted by Nunnally, Stevens, and Hall (1965).

The present experiment was designed to investigate whether repeated ratings of stimuli in the Nunnally paradigm, described above, partially account for conditioning effects by increasing demand characteristics or similar mechanisms. The effect of ratings can be determined by varying the number of ratings and by varying the stimuli rated from one learning occasion to the next. These controls provide a cross-check on the hypothesis that repeated ratings may have been a dominant variable in the Nunnally, Stevens, and Hall (1965) study. As such, the rating procedures might have spuriously inflated the degree of preference attributable to conditioning of meaning (to be described in a later section in greater detail) or it might have increased only Ss' knowledge of the demands of the situation thereby resulting in pseudo conditioning. If the linear increase in positive verbal evaluation which those authors obtained over days was a product of mediated awareness through ratings rather than continuous strengthening through reinforcement, the proposed design should provide the necessary demonstration.

In general, if preferences increase as a function of the number of evaluations made in a situation, clearly, multiple ratings is a factor which must not be ignored in such studies. If preference with only one rating, at the end of the experiment, is comparable to the final rating for Ss who make a number of ratings, then the conclusion that conditioning has been the dominant factor is a valid one. However, this would still not preclude the possibility that learned demand characteristics (through instructions, subtle cues given by E, etc.) have influenced that rating (Orne, 1962). For purposes of assessing such contributions, efficient post-experimental measures of S's hypotheses and intents are invaluable (Dulany, 1962).

The findings should also be valuable in the evaluation of other semantic conditioning experiments. Findings such as those predicted for this investigation would have implications for current theories which feature versions of the classical conditioning paradigm for the conditioning of meaning. The study is not meant as a challenge to all demonstrations of "learning without awareness." Many kinds of behavior occur without awareness (Adams, 1957). But the establishment of conditions under which cognitive hypotheses can be unequivocally demonstrated remains an interesting and important problem for investigation.

CHAPTER II

REVIEW OF LITERATURE

This review contains studies which are relevant to what is commonly termed "conditioning of meaning." Its emphasis is on the S's perception of the demands of an experiment and the relationship of such perceptions to learning without awareness.

The role of "awareness" in the experimental study of learning has a long history, dating back to Thorndike's time, although interest in it has enjoyed a great expansion during the past decade. Various aspects of behavior without awareness such as posthypnotic suggestion and automatic writing were of greater interest to early investigators than were the effects of awareness upon learning per se as indicated in reviews by Coover (1917) and Miller (1942).

Automatic Strengthening and Cognitive Influences on Conditioning Involving Verbal Materials

The Law of Effect

Probably the most polemical feature of Thorndike's (1932) law of effect is the proposition that the action of rewards is direct, automatic, and independent of volitional perceptions and cognitions on the part of the learner. In recent years the arguments have been renewed through increased interest in (a) the relationship between reinforcement and intent to learn and (b) learning without awareness (Postman & Sassenrath, 1961). Thorndike recognized the importance of testing the premise that rewards and punishments can influence S-R connections even while the S remains unaware of what he is learning. He performed a number of experiments directed at this question and found support for the influence of aftereffects when S does not understand what is happening in the situation (Thorndike, 1932).

Thorndike (1935) gave five reasons for believing that such automatic action of reinforcement did take place: (a) S seldom reveals during questioning that he has deliberately considered the alternatives; (b) the connection has a strength of 1.00 before Ss have confidence in its correctness; (c) rewards can be effective even when S does not know why he is being rewarded; (d) deliberate discrimination among possible responses should cause S to avoid incorrect responses about as often as it leads to repetition of correct responses; (e) a cognitive interpretation does not account for the spread of effect from rewarded connections to neighboring punished ones (Postman & Sassenrath, 1961).

One experiment (Thorndike, 1932) featured cards with four lines of identical length. S's task was to judge which of the four lines was longest. Other identifiable characteristics (e.g., ink blots) on the cards allowed S to recall specific cards and to repeat the rewarded answer for that card. "Unaware" Ss increased in the number of "correct" responses they gave. The criterion used for awareness in the experiment is similar to verbal reports discussed in a later section.

A further pursuit of the topic followed in the well known Thorndike and Rock (1934) study in which a free association method was used. Syntagmatic (sequential) or rote responses rather than paradigmatic (same form class) denotation or defining responses were reinforced. The investigators assumed that insight into the principle governing the reinforcement would produce a sudden decrement in errors. Since improvement was gradual instead of sudden it was proposed that the strengthening of responses had been automatic.

Thorndike's conclusion does not appear to be valid in view of a slightly modified replication of the Thorndike and Rock experiment by Irwin, Kaufman, Prior, and Weaver (1934). In that investigation each _S_ was taught the principle governing the correct response at some time during the session. Improvement was gradual both before and after the insertion of the instruction. Thus, there was some question whether Thorndike and Rock had presented conclusive evidence for learning without awareness.

The casual reader of Thorndike might conclude that a theory of automatic strengthening of the S-R bond through the law of effect leaves no room for conscious processes. On the contrary, his position was that persons are often aware of the contingency between responses and rewards and that the effects of learning without awareness which he obtained in his laboratory were, in general, slight. "Unconscious" learning was viewed as relatively inefficient and inconsistent and conscious awareness of this relationship as more effective. "We should, of course, make the situation identifiable and the response available when it is practicable to do so. Unconscious learning is relatively undependable and slow" (Thorndike, 1935, p. 70). Furthermore, he held that aftereffects of responses have not only a strengthening, but also an informative influence involving ideational content (Farber, 1963).

Recent Investigations of Learning Without Awareness

A number of studies and comments related to Thorndike's work on learning without awareness followed these early studies. A controversy over continuity versus noncontinuity learning is traceable to Lashley (1929) who insisted that practice preceding an association is irrelevant to its formation. In this paradigm the presolution period represents the testing of several alternative and independent hypotheses until a successful one is identified for solving the problem. These attempted solution hypotheses are based upon cognitive belief that particular cues selected lead to the demanded significate. The proponents of this position proposed that while an *S* is responding on the basis of an incorrect hypothesis, he learns nothing at all about the relevant cues. Spence (1945) along with the majority of the investigators on this topic favored the continuity explanation in which each reinforced response to the positive stimulus is said to strengthen the tendency to respond to it in incremental fashion. This interpretation of the law of effect all but neglected the role of cognitive hypotheses in facilitating learning. As the current literature attests, both the initial continuity and noncontinuity positions were too extreme since verbal learning is clearly shaped by associative and cognitive processes.

In such experiments involving humans, the S usually seeks the reinforcement and at least partially understands its relevance.

However, Wallach and Henle (1941, 1942) asked whether rewards increase the probability of the previous response being repeated even when the S does not understand the relevance of the reinforcements. This would virtually eliminate the possibility of deliberate rehearsal and selection of the rewarded response. They used "Right" and "Wrong" to reinforce Ss' guesses of numbers as responses to a series of words. It was presented as "a study of E.S.P." in which S was to name the number that another person had for a particular word. It was implied that a response called "Right" on one trial might or might not be correct on the next trial. No support for the theory of automatic strengthening of responses was found under these conditions.

More recently, Postman and his associates as well as others have contributed considerable insight into the nature of the law of effect. Postman and Adams (1954, 1955), modifying the Wallach and Henle E.S.P. procedure to eliminate a methodological flaw, found that rewarded responses were repeated and recalled significantly more often than punished ones.

Postman and Jarrett (1952) repeated the Thorndike and Rock (1934) experiment discussed above but modified the procedure in accord with that used by Irwin, Kaufman, Prior, and Weaver (1934). After each block of trials S stated the basis on which he was responding. These frequent inquiries may well have caused S to actively seek a principle. Again, the results of the experiment showed little evidence for learning without awareness. Although the principle involved might be verbalized by *S*, its application was difficult enough that skill in applying the principle improved steadily and gradually even after the principle was understood.

Using the same paradigm but making the principle simpler and easier to apply, Philbrick and Postman (1955) found a sudden increase in correct response following verbalization of the principle. The *S*s who were unsuccessful at verbalizing the principle of correct response, the number of letters in the stimulus word minus one, also showed a significant amount of learning. However, their performance was consistently poorer than that of the *S*s who verbalized a correct hypothesis. Thus, the simple principle facilitated learning without awareness as well as improvement resulting from awareness. Using the same task as that used by Philbrick and Postman, Hirsh (1957) also found support for an automatic affereffect on S-R connections independent of *S*s understanding.

Postman and Sassenrath (1961) defend Thorndike's work on the automatic action of reinforcers against Adams' (1957) criticisms. In his review Adams casts serious doubt on most of the existing evidence for learning without awareness. Postman and Sassenrath see no sharp dividing line between learning with or without awareness. "Since verbalization often occurs after a period of systematic improvement, *verbalization of a principle may be considered at the same time a result of past improvement and a condition of further improvement*" (Postman & Sassenrath, 1961, p. 124). Or as Verplanck (1956) has put it, verbalization can be thought of as both an independent and a dependent variable. As an independent variable, it is an antecedent condition of performance; as a dependent variable, it is a function of the conditions of reinforcement.

The study by Hirsh (1957), cited earlier, used three types of consequences, reward, punishment, and reward plus punishment. The rank order of the conditions with respect to sudden "insights" or verbalizations not preceded by any evidence of improvement was: (a) reward plus punishment, (b) reward, and (c) punishment. Postman and Sassenrath (1961) indicate on the basis of Hirsh's experiment that the less favorable the conditions of reinforcement (e.g., punishment alone), the more likely is verbalization to be preceded by a period of improvement without awareness. On the other hand, when conditions of reinforcement are highly favorable (e.g., reward plus punishment), improvement in performance and verbalization of the principle may coincide in time. Accordingly, verbalization can be a special class of responses which is strengthened as a function of the conditions of reinforcement, and it can also be a condition of improvement.

Despite this 1961 statement by Postman and Sassenrath a number of reputable authors were in close agreement with Adams' (1957) position. Dulany (1961), for one, concluded that ". . . a theory of automatic strengthening by aftereffects apparently does not account for these findings without auxiliary assumptions or radical augmentation" (Dulany, 1961, p. 261), a view strongly supported by Spielberger (1962).

Awareness As It Relates to Verbal Conditioning

Farber (1963) suggests that three classes of variables are relevant to understanding the nature, causes, and effects of awareness: motivational variables as they affect perception and performance; mediated generalization and discrimination processes; and verbal operant conditioning.

Most of the recent interest in learning without awareness (LWOA) stems from the last-mentioned area which can be traced to an experiment by Greenspoon (1950) in which Ss' verbalizations could be modified by saying "mmm-hmm" at appropriate times. Thus, the frequency of either plural or nonplural responses was increased through E's verbal reinforcement of words in one class or the other as they were elicited by S (see, for example, Greenspoon, 1955). Data from aware Ss were eliminated from this study.

Sidowski (1954) used a light as a reinforcer in the same paradigm employed by Greenspoon and found that learning occurred without knowledge of the response-reinforcement contingency. Cohen, Kalish, Thurston, and Cohen (1954) reported that no awareness was revealed by questioning of Ss in their experiment. On each trial Ss chose one of six personal pronouns to form a sentence with a verb. E said "Good" when "I" or "We" was chosen by S. They found a linear increase in "I" and "We" responding for reinforcement as compared to a nonreinforced control group.

A field study by Verplanck (1955) similarly revealed a significant change in behavior without evidence of awareness. Casual conversations were found to support they hypothesis of ". . . increase of opinion statements through agreement." In a footnote, Verplanck reports that he also demonstrated the effect while describing the investigation to a colleague. The colleague reinforced him through agreement, and, like others in the experiment, Verplanck was unaware that he had been an S until informed.

Further evidence for LWOA followed in the next year with the free association work of Eriksen and Kuethe (1956). Ss were shocked for giving certain selected free associations and classified as high or low aware via questions at the conclusion. A significant decrement in the number of formerly punished responses given during chained associations was found for both groups even when Ss knew that shock would no longer be given. No reliable difference was found between these two groups for learning to avoid the shock; however, the reaction times for the two groups were clearly different with the aware Ss needing more time to mediate, apply a rule, and make a response.

In a study by Essman (1957) Ss increased their tendency to assign nonsense shapes to "Human" or "Anatomical" categories when E said "good." Although eight of ten Ss were aware that E was saying something, the author reports that the statements were unrelated to the purpose of the experiment.

On the basis of the above review, it appears that most of the early studies during the 1950-1958 era fail to support the notion that Ss become aware of the reinforcement contingency in verbal conditioning. Fewer than five percent of Ss in studies reviewed by Krasner (1958) had been classified by Es as aware. In some cases the investigators (e.g., Wilson & Verplanck, 1956; Finley & Staats,

1967) have concluded that even though Ss might accurately identify the reinforcement and the correct response class, it is not a significant problem from a performance standpoint. Nevertheless, the problem is an interesting one to psychologists, and continuing explorations have provided evidence against LWOA. In some studies aware Ss, compared to unaware Ss, show a decidedly greater increase in the frequency of usage of reinforced word classes (Farber, 1963; Eriksen, 1960). In addition, a number of experiments have demonstrated that any increment in correct responses was entirely due to Ss classified as aware (Tatz, 1960; Gerstein, 1961; Levin, 1961; Spielberger, 1962; Weinstein, 1963; Spielberger, Berger, & Howard, 1963; Eriksen & Doroz, 1963; Lanyon, 1964; Katkin, Risk, & Spielberger, 1966).

Modified verbal conditioning studies by Di Vesta and Blake (1959) and by Di Vesta and Walls (1967a), as well as a replication of the former study by Janke (1962) indicate than an individual's "set" to look for some principle for responding is an important consideration in any LWOA paradigm. Postman (1962) says that a set to find a principle had a beneficial effect on performance in the Di Vesta and Blake (1959) study. Those authors interpreted their findings as providing evidence that "instructions permit verbalizations which mediate adaptations to the learning task by the formulation of hypotheses" (Di Vesta & Blake, 1959, p. 67). The relationship of this study to later discussions of mediation and hypothesis formation should be borne in mind.

The earlier studies, related to this topic, have been particularly vulnerable to attack. Lacey and Smith (1954) have shown that the well-known Diven (1937) study on unconscious conditioning of anxiety did not logically demonstrate conditioning without awareness. However, even with their more sophisticated design, they did obtain some evidence of unconscious conditioning. But again, Chatterjee and Eriksen (1960) have demonstrated that Lacey and Smith's evidence may have arisen from an artifactual method of computing scores. These latter investigators, using a more careful assessment held that autonomic conditioning was no more precise nor specific than the *Ss*' verbalizations (Eriksen, 1960).

In an attempted replication of the oft-cited Rees and Israel (1935) study, Davis and Hess (1962) found close relation between awareness and anagram solving efficiency, contrary to one of the main implications of the Rees and Israel study. Taken together, the experiments suggest that a mechanism of cognitive control may be widespread in verbal learning and conditioning. Thus, *Ss* may hypothesize what is expected of them and instruct themselves to act accordingly (Dulany, 1961).

Criteria for Awareness

LWOA studies have been plagued with the problem of adequate operational criteria for awareness. The most often used definition of awareness and unawareness is in terms of verbal report. Awareness is generally equated with the ability to verbalize the response-reinforcement contingency and unawareness with the inability or lack of verbalization. Eriksen (1960) comments that the popularity of a definition of awareness in terms of verbal report is understandable. It creates the illusion of operational precision and perhaps masks the fact that there may be more important characteristics that are avoided by such a definition.

The discrepancy between inferences from the early and later investigations in verbal conditioning and learning appears to be largely a function of the method used by E to assess awareness. The later works typically employ systematic and sophisticated inquiry to determine S's subjective perception of his task and role in the experimental setting, using intensive postexperimental inteviews as the primary source of information.

Groups classified as differing in awareness have been compared on many variables and in several instances given the status of experimental groups with awareness being treated as an independent and antecedent variable, e.g., the study by Eriksen and Kuethe (1956) cited above. When awareness is inferred from response variables, e.g., Thorndike and Rock (1934), it is inconclusive whether awareness causes the responses or is only correlated with them (Martin and Dean, 1965).

It has been suggested (Dean and Hiesinger, 1964) that perhaps Greenspoon's brief interview, which followed a lengthy extinction period, may not have detected all Ss who were aware. Because of his long (25 minute) extinction period, some Ss may have forgotten the correct response by interview time. Wong, Harrison, and Stopper (1966) found no evidence of awareness with Dixon and Oakes' (1965) interview procedure, which was relatively superficial, but with a detailed interview conditioning was seen to occur only with awareness. The importance of intensive interviewing procedures immediately following acquisition trials is reported by several authors (Weiss, Krasner, & Ullman, 1963; Krieckhaus & Eriksen, 1960; Levin, 1961). Spielberger and DeNike (1963) agree with Farber's (1963) statement, which indicates, "If it were not for an unreasonable prejudice against questioning Ss, psychologists would long ago have discovered what everyone else already knows, that one can usually understand a person's behavior much better if one tries to find out what he thought of the experiment and decided to do about it" (p. 187).

It is proposed by Postman and Sassenrath (1961) that if S cannot state the conditions of reinforcement which systematically influence his behavior, then he cannot discriminate the information transmitted by the reinforcement. "Hence such learning must reflect the automatic action of aftereffects and not the use of information which can be rehearsed and recalled at the time of the test" (Postman & Sassenrath, 1961, p. 111). It appears to the present reviewer, in view of the limited verbal facility of young children and some adults as well as possible alternative methods for assessing awareness, that Postman and Sassenrath's (1961) statement was a bit hasty. For example, Grings (1965) reports that classical conditioning can be as reliably obtained with retardates as with college students, and it is not possible to obtain verbalizations that allow E to determine whether Ss were aware or not. Spielberger (1962) indicates that a large proportion of the positive evidence for verbal conditioning without awareness has come from studies involving psychiatric patients (e.g., Buss & Gerjuoy, 1958; Cohen, Kalish, Thurston, & Cohen, 1954; Leventhal, 1959; Taffel, 1955). Since such patients are often relatively inarticulate with inadequate vocabularies and vague conceptual categories, the relationship between awareness and learning may be difficult to evaluate through verbalization. As Kimble (1962) notes, it is probably inappropriate to even discuss the matter of verbalization for ". . . young children, certain of the very old, many brain damaged individuals and most mental defectives . . ." (Kimble, 1962, p. 29).

Several other methods of assessing awareness and variables relevant to such assessment have been suggested by various authors. Increments in performance first occurred on the trial block on which _S_s first recorded the correct hypothesis on written "thoughts about the experiment" in an experiment by DeNike (1964). Shupe (1966) suggests that a postexperimental multiple choice questionnaire offers advantages of objectivity and control over the more typical interview. A "visual recognition threshold task" is proposed as yielding reliable and satisfactory data on degree of awareness by Klein and Wiener (1966). This recognition procedure utilizes a booklet consisting of 25 successive carbon copies arranged from least clear to the clearest copy; a technique similar to that described by Cowen and Beier (1950). Mandler and Kaplan (1956) questioned _S_s as have most _E_s but then scaled the replies on an 11 point level of awareness scale. Full awareness of the reinforcing contingency was assigned a score of one with 11 representing complete lack of awareness. An obtained range of three to 11 and a mean rating of six indicated that the average *S* verbalized some secondary or correlated hypothesis "that there might have been some unknown relationship between his associations and the experimenter's behavior" (Mandler & Kaplan, 1956, p. 582). A different procedure, (Orne, 1959) involves a control group which goes through the inquiry procedure as though they had actually been run in the experiment, without, however, having actually participated in data production. Simulating *S*s have also been used, particularly in hypnosis (Orne & Evans, 1966).

Thus, it can be seen that a number of variations in assessment of awareness have been proposed and utilized. However, it appears that the probability of obtaining awareness reports in verbal discriminations is at least partially a function of the class of discriminative stimuli. Kanfer and McBrearty (1961) employed the Taffel-type (1955) verbal conditioning task in which *S* constructs a sentence using either a hostile or a neutral word. With reference to Skinnerian theory they call the rewarded response class (hostile) an S^D and the never reinforced class (neutral) the S^Δ. But they emphasize that verbal conditioning is not a simple case of operant conditioning. They found that the greater the distinctiveness between the S^D and S^Δ the higher the proportion of aware *S*s. Thus, the probability of obtaining awareness reports in verbal discrimination is partially a function of the class of discriminative stimuli.

Semantic Conditioning of Meaning

On the basis of work by Spielberger (1965) and many other investigators who have been able to account for a considerable portion of the variance by the degree of awareness measured in verbal conditioning situations, it appears reasonable to conclude that a simple extension of operant conditioning principles developed at the animal level cannot adequately account for details of verbal conditioning (Maltzman, 1966). But what of the classical conditioning paradigm in which one would expect the responses to be less volitional and more automatic? The classical theory of automatic strengthening, unqualified by further assumptions, clearly implies that learning should occur in the absence of either a correct or a correlated hypothesis. This has been demonstrated with human *S*s a number of times.

For example, Newhall and Sears (1933) report that a "sophisticated" *S* was surprised to learn that he had made conditioned finger responses when he thought his finger had not moved during any of the test trials of the session. Likewise, *S*s used by Hefferline, Keenan, and Harford (1959) remained ignorant of an invisible small thumb twitch which had been conditioned to terminate or postpone aversive noise. Other *S*s, informed of the effective response, could not produce it deliberatley in a size small enough to qualify for reinforcement. But as Spence, Homzie, and Rutledge (1964) point out, even in a highly controlled situation involving human eyelid conditioning where a voluntary blink can be distinguished from a conditioned one, avoidance of voluntary factors by masking the true purpose of the experiment is highly desirable.

Relatively few studies have actually used verbal materials for the classical conditioning model (Bulgarella, 1966; Golin, 1964; Razran, 1949a; Razran, 1949b). The Staatses (1957) reasoned that if a nonsense syllable is immediately (one second in this investigation) followed by a meaningful word, the meaning of the word would be conditioned to the nonsense syllable. However, if the nonsense syllable is paired with the same word each time, and the nonsense syllable did, in fact, come to elicit the same meaning as the word, it might be accounted for by direct association. The authors determined to eliminate such a possibility by pairing the syllable on each trial with a different word. It was necessary that each of these words reflect the same connotative meaning. For example, the words pretty, sweet, and healthy all have positive evaluative connotation, and yet are quite different denotatively. Thus, the nonsense syllable is considered to be a conditioned stimulus (CS) and the words to be unconditioned stimuli (UCS). The semantic differential, an instrument developed by Osgood (1952), was used as the method for assessing the meaning built into the syllables. There was significant evidence that meaning responses had been conditioned to the nonsense syllables.

In addition to nonsense syllables, those authors have conditioned meaning to proper names (Staats & Staats, 1958) and meaningful words (Staats, Staats, & Biggs, 1958). They also found semantic generalization of conditioned meaningful words to synonyms (Staats & Staats, 1959).

Pollio (1963) employed the Staats' procedure involving three nonsense syllables as CS's and words as UCS's. Following conditioning, Ss gave word associations to the CS syllables. These associations were then rated on the semantic differential by other Ss; these Ss produced ratings similar to those for the original UCS's. Although the Staats' paradigm results have not always been easy to replicate, this study by Pollio strongly supports their findings.

Using the technique developed by Staats and Staats, Cohen (1964) paired nonsense syllables (CS) with words (UCS) of high and low evaluative meaning. After the conditioning trials, Ss rated the syllable on a 7-point semantic differential scale whose end points were labeled Pleasant and Unpleasant. Contrary to the results of Staats and Staats, no evidence for LWOA was found.

Research in Maltzman's (1966) laboratory on semantic conditioning and generalization uniformly shows that aware Ss are reliably superior to unaware Ss and that individuals who do not verbalize the CS-UCS contingency do not manifest performance which can be interpreted as successful conditioning. However, Maltzman also finds that institutionalized schizophrenics, like the retardates mentioned earlier, may show reliable semantic conditioning and generalization even though, because of their deterioration, it is not possible to obtain a coherent verbal report. Although there is no verbal evidence of whether or not Ss are aware, they show semantic conditioning.

Thus, we have seen that from the time of Thorndike investigators have taken sides in the awareness-automatic strengthening controversy without resolution. It appears that both processes must be carefully considered in any conditioning of meaning experiment in order to better understand the effects of the manipulated variables upon the dependent measures.

The Influence of Mediated Experimental Demand Perceptions on Theory and Experimentation

The Relation of Mediation to Awareness

One commonly finds "mediation" invoked as the vehicle via which Ss recall and generalize the CS-UCS contingency (Kimble, 1962). In order to better understand the term mediation as it applies to semantic materials a brief discussion of Osgood's (Osgood, Suci, & Tannenbaum, 1957) theory of meaning is helpful. The name of an object is the "sign." This sign is capable of evoking a representational mediating response which is some fractional part of the behavior that would normally occur to the object itself. Mediating cues are produced by this mediational response that would not have occurred if the sign (word) had never been associated with the object. Because of these representational components ($r_m \rightarrow s_m$) nonsense figures, syllables, and other stimuli with no inherent meaning may take on some of the meaning of the sign through conditioning. When these higher order neutral stimuli take on such meaning they are termed "assigns."

The relationships are illustrated by Osgood, Suci, and Tannenbaum (1957) in Figure 2. The $r_m \rightarrow s_m$ is part of the same behavior (R_T) produced by the significate or object or concept $(\dot{S})$. The word or name is a sign $(\boxed{S})$ of the object or concept $(\dot{S})$ rather than a sign of any of a multitude of other things. In Part B of Figure 2, Osgood et al. diagram the development of the assign.

As can be seen in that figure, assigns (/S/) have their meanings "assigned" to them through association with other signs rather than through direct association with the objects signified. Thus, most of the signs that we use are really assigns. Concepts such as "good" and "awful" take on the representative connotation of any number of signs.

The conditioning of meaning experiment is closely related to Osgood's theoretical framework. In an important study of the mediation of evaluative or attitudinal responses, Eisman (1955) investigated selective preference in children. It was found that the block to which a rewarded color name had been attached was preferred significantly above a theoretical baseline.

Osipow (1960) demonstrated that color names associated with either positive or negative evaluative words led to similar changes in preference for nonsense figures associated with those color names. In a more closely controlled experiment (Di Vesta, 1962) preferences for a nonsense figure associated with a color name were changed. Both reinforcing a neutral color name and attaching a neutral color name to a positive evaluative sign in the first stage of the mediation process led to significant changes in preference for the manipulated figure.

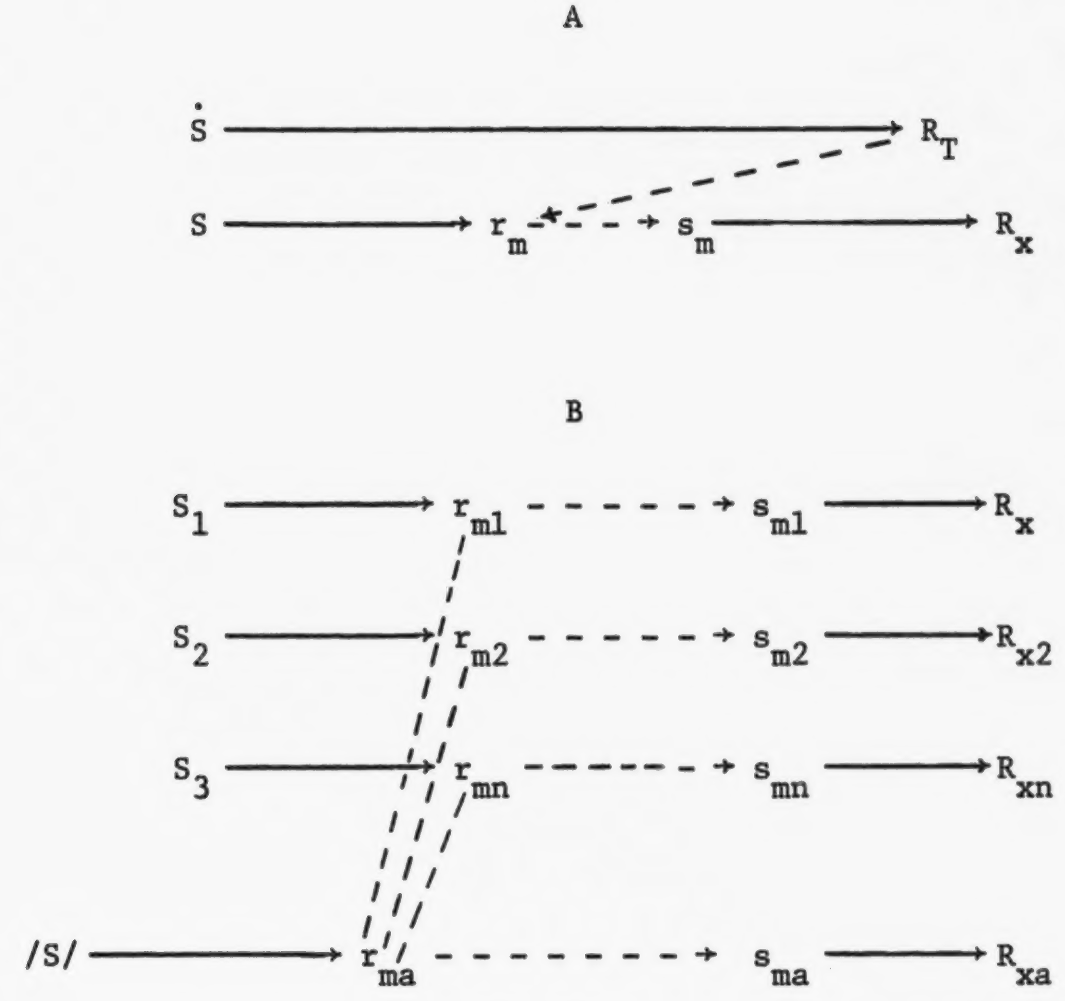

A
Ṡ
R_T
S
r_m
s_m
R_x
B
S_1
r_m1
s_m1
R_x
S_2
r_m2
s_m2
R_x2
S_3
r_mn
s_mn
R_xn
/S/
r_ma
s_ma
R_xa

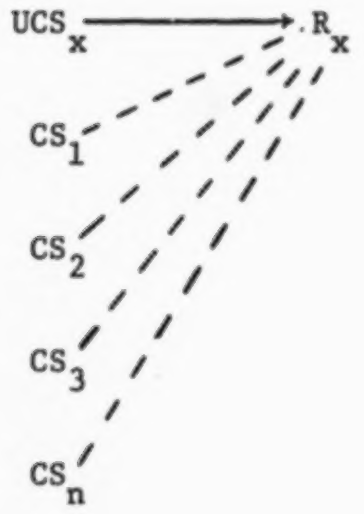
Figure 2. (A) Development of a sign; (B) Development of an assign (Osgood, Suci, & Tannenbaum, 1957).

Another mediation study from the same laboratory (Di Vesta & Stover, 1962) involving elementary school children consisted of three phases. First, assigns were developed by conditioning signs with evaluative meaning to neutral nonsense syllables. Next, the assign was associated with a neutral nonsense figure. In the third phase the nonsense figure was rated by Ss on evaluative scales of the semantic differential. A second experiment enlarged upon and supported the general findings that ratings of the figures labeled with the conditioned assigns corresponded with the evaluative meaning of signs associated with those assigns. Thus, we see that in Osgood's terminology, when an assign is associated with evaluative signs, the meaning will depend upon the nature of the mediating response evoked by these signs.

However, a number of other representations and proposals of the nature of mediation besides Osgood's "detachable" fractional response have been expressed (Cofer & Foley, 1942; Mowrer, 1954; Spiker, 1956). For Cofer & Foley (1942) mediated or semantic generalization depended upon preexperimental formation of conditioned responses or associations. Figure 3 suggests the way these authors viewed semantic generalization in 1942.

S comes to the experiment with the tendency to make R_x to CS_1, CS_2, CS_3, . . . CS_n or else this may be built in during a pretraining session. Then S is experimentally conditioned by reinforcing CS_1 with UCS_y, setting up a conditioned response, R_y. Generalization is now presumably found to CS_2 . . . CS_n. Invoking the theoretical analyses of Hull (1938), Miller and Dollard (1941)

Preexperimental CR's


UCS$_x$
R$_x$
CS$_1$
CS$_2$
CS$_3$
CS$_n$

Experimental CR's for Testing Mediated Generalization

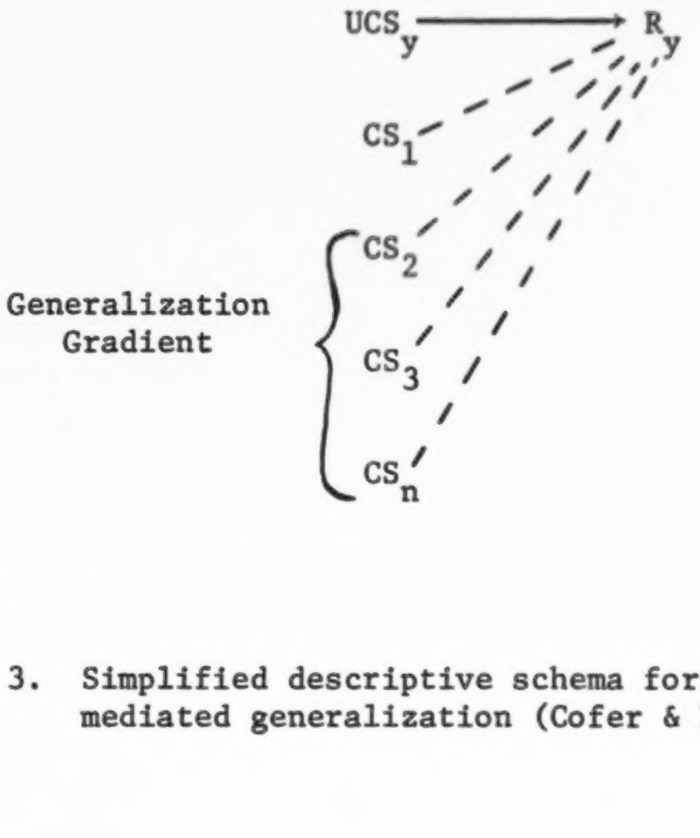

UCS$_y$
R$_y$
CS$_1$
CS$_2$
CS$_3$
CS$_n$
Generalization Gradient

Figure 3. Simplified descriptive schema for the development of mediated generalization (Cofer & Foley, 1942).

and others, Cofer and Foley (1942) indicate that upon the pre-experimental (or pretraining) conditioning of $CS_1 \ldots CS_n$ to R_x, each of these stimuli is also conditioned to an implicit, fractional, kinesthetic r_x. When $CS_2 \ldots CS_n$ is presented alone for testing, it will evoke r_x, whose s_x will in turn elicit R_y.

For Mowrer (1954) a sentence is a conditioning device. He posited a "mediating reaction" which the subject of a sentence (as well as the referent of the subject) produces. By assuming that the meaning reaction produced by the subject of a sentence is conditioned to the mediating response rather than directly to the sentence subject, he accounts for the phenomenon of semantic generalization. Using meaning as an implicit mediating response, he suggests that communication takes place when the meaning response which has been elicited by the predicate is conditioned to the subject of the sentence.

A discussion by Di Vesta (1962) is based upon a detailed formulation of mediated generalization by Spiker (1956), in which S experiences two stimuli, a figure (S_F) and a color (S_C). These are presented contiguously to S and he is trained to make a common verbal response (R_x) to both of them; for example, he calls both "green." Then the S learns to respond with "good" (R_y) to the color "green" (S_C). While this new $S_C - R_y$ connection is being learned, s_x (the stimulus produced by response R_x) also becomes able to elicit R_y. Thus, S_F will gain in subjective evaluation. The explanation then is that the figure (S_F) elicits the verbal response (R_x); R_x then elicits s_x, and s_x elicits R_y response. These discussions of mediation that have been presented here, differ little in theoretical structure, but their general representation should be borne in mind for later explication.

Other papers by Di Vesta (1966), Carroll (1964), and Miron and Osgood (1966) discuss mediating processes and the way in which meaning is acquired. Carroll suggests that through general contact with adult speech the child learns the denotative meaning of the criterial attributes of signs or words. However, there are also noncriterial attributes which are irrelevant to denotation that come from individual experience. Through classical conditioning, these responses are associated with linguistic symbols and correspond to connotative meanings. To the extent that people have similar experiences with a certain referent or language symbol, they will express the same connotation for it. For example, pretty and beautiful are both connotatively "good" for most individuals while ugly and messy have a "bad" connotation. Words such a pretty and ugly have strong denotative value as well as connotative. Words like good and bad are evaluatively, more general in nature, and contain more of the affective element without denoting anything in particular. It is a relative distinction that has been made most systematically through the semantic differential.

Both Carroll (1964) and Miron and Osgood (1966) believe that an individual's evaluative system develops through the person's interaction with the environment. Initially, it is induced directly by the stimulus attributes. Later, fractional representations are induced by signs of the original stimulus. Consequently, we can think of evaluation of an originally neutral stimulus as eventually coming under control of originally inadequate signs produced by those persons with whom we interact. Measurements by the semantic differential on the evaluation scales or some similar rating device can provide an index of the satisfaction value of the stimulus to the person.

White (1966) has done considerable work with children from age three to eight. During this period certain shifts in learning and preference appear to be taking place (also see Kendler, Kendler, & Wells, 1960). Verbal mediation appears and has an important effect on two stage learning processes. In addition a shift in position preference as well as one for touching versus seeing is going on about first or second grade. However, Di Vesta (1966) and Miron and Osgood (1966) conclude that children's semantic structures, beyond the second grade, at least, are very much the same as those found in adolescents and adults. Such mechanisms are implicit in the evaluative structure of even the seven year old. The Evaluation dimension, as opposed to the Activity and Potency dimensions (Di Vesta, 1966) is more heavily weighted both for children in the second or third grades and for older children.

Children of this age and younger have difficulty verbalizing their perceptions of experimental contingencies, and most of the debate concerning the assessment of learning without awareness has centered upon the validity of Ss' verbalizations as a criterion of awareness. Related to this difficulty is the possibility that different levels of awareness exist and that verbalizable contingencies need not be present to demonstrate conditioning.

As Postman and Sassenrath (1961) put it, "It is an outmoded and unnecessary assumption that the modification of behavior must be preceded by a correct understanding of the environmental contingencies" (Postman & Sassenrath, 1961, p. 134). They mean by this statement that whether or not the action of rewards is always or never automatic is not a profitable question. Verbalizations may be considered to be mediators which act to increase the probability of appropriate responding. Unverbalized mediators, on the other hand, are hypothesized to be involved in producing LWOA.

Previous studies have consistently demonstrated that mediation can occur in the absence of awareness (Bugelski & Scharlock, 1952; Russell & Storms, 1955; Mednick & Freedman, 1960; Horton & Kjeldergaard, 1961). In a three-stage mediation paradigm Horton (1964) deliberately manipulted conditions to induce varying amounts of awareness. He reasoned that the aware _S_s, particularly in view of the findings reported in verbal conditioning literature (e.g., Dulaney, 1962; Spielberger, 1962), should generate maximum mediational effects within the "limits" of a particular mediation paradigm.

Horton's findings were in line with this prediction; that is, when experimental conditions provide cues that facilitate a knowledge of the operation and purposes of the experiment, the unaware _S_ is likely to be a particularly poor performer. This lack of performance has been attributed to task set, strategy, or a variety of other variables which hinder some _S_s from attending to task relevant cues. However, the work of Horton and others have demonstrated that when conditions are not particularly conducive to awareness, many of the unaware *S*s can perform quite well. They found that both meaningfulness of the common element and learning ability of *S*s were significantly related to awareness in this mediation situation.

The Perception of Experimental Demands

Orne (1962), in commenting about procedures for assessing awareness and conscious verbal mediation, says that such inquiry techniques as those discussed earlier clearly have their own "demand characteristics." In this view, *S* is considered to be both an active and productive respondent. Thus, his behavior in any experimental situation will be determined by two sets of variables (a) those traditionally defined as experimental variables, and (b) his cognition of the way in which he should respond. He usually will try to be a "good *S*" or behave in a manner that is beneficial to the *E*'s experimental hypothesis. Regardless of whether or not *S*'s behavioral hypothesis is correct, his responses will be influenced. In a similar vein, Di Vesta (1962) points out that although mediated generalization may be one basis of the development of preferences, it is desirable to separate supposed secondary reinforcing properties that have accrued through reinforcement from conscious tries to be right.

Tolman (1932) believed that *S* forms definite cognitions about what to expect and acts upon these expectancies. *S* must experience a reward "one or more times" before his "behavior can change so as to become appropriate to it" (p. 71). He "must have a chance to build up a 'cognitive expectation' for this new reward" (Tolman, 1932, p. 71) with its given demands.

In the excellent Journal of Personality monograph on awareness which appeared in 1962, Verplanck (1962) concurs that sooner or later the S will start "making hypotheses" about features of the experiment. And, while verbal behavior cannot be expected to bear a one-to-one relationship with concepts of "hypotheses," "mediators," or perceived demands, there can be no dispute that these are involved in what we mean by "awareness."

The demand characteristic perceived in any particular experiment will vary with the sophistication and previous experience of the S. Shepherd (1957) found this to be the case in a discrimination learning-to-learn experiment with children. She used upper and lower case Greek letters randomly arranged in six pairs and informed the children that the "pictures would tell them which box to pick." The Ss reached asymptote on the second problem (six trials with each problem). Shepherd as well as Harlow (1949), the formal originator of the learning-to-learn principle, found that children as opposed to monkeys are fast at "catching on," and learning sets in some cases are formed with relative ease. Other studies have supported the findings that human learning and hypothesis testing situations which extend over a long period of time are often influenced by learning set formation (see for example, Di Vesta & Walls, 1967a; 1967b; 1967c).

A Theoretical Structure for Assessing Perceived Demands

If, as the evidence strongly demonstrates, humans do form hypotheses mediated by self instructional sets about the demand characteristics in an experiment, then precise standardized measures of these phenomena as they relate to awareness are necessary. Dulany (1962) believes that such concepts as _S_'s hypotheses enter into lawful relationships no less fundamental than any other. Self-generated hypotheses may influence behavior just as a direct experimental manipulation can. When we treat _Ss_' reports of hypotheses and actions in an experiment as unimportant we may have been discarding the best part of the experimental data. With an awareness of perceived demands encompassed by theory, an alternative strategy becomes more appealing.

Such a strategy is suggested by Dulany (1962). He stated that, ". . . a human _S_ does what he thinks he is supposed to do if he wants to . . ." (Dulany, 1962, p. 109). A theoretical network was then proposed consisting of a "self-instructional set" mediating the responses. This set is divided into three parts, as follows:

> A _Reinforcement_ _Hypothesis_ in which _S_ thinks that the reinforcement
>
> (a) signifies that a preceding response is correct;
>
> (b) merely follows a certain response;
>
> (c) only occurred; or,
>
> (d) did not occur.
>
> A _Behavioral_ _Hypothesis_ in which _S_'s hypothesis is that a particular response or response class is correct and is what _E_ expects _S_ to do or choose.

> A Behavioral Intention which is a "volitional" intention to select certain members of the response class.

The postexperimental questions that Dulany (1962) uses to ascertain the nature of these three concepts for a particular S more fully define the terms. They are as follows:

1. a. Did you notice whether or not I said anything during the experiment? b. (What?)
2. a. Did you come to think it was random or did it follow anything in particular that you did? b. (What?)
3. a. Did you come to think there was or wasn't any purpose or significance to the "Good" in this experiment? b. (What?)
4. a. Did you come to think that there was anything you were supposed to say, or not say on each trial in order to be correct--something the experimenter wanted you to say or not say? b. (What?) c. (Did you come to think there was or wasn't any kind of correct response?)
5. a. How did you go about making up your sentences? b. How did you go about selecting a word from the bottom of the card? c. (Would you say that you did or didn't try to use any particular words or kinds of sentences?) d. (What?)

Questions 1, 2, and 3 are directed to the S's Reinforcement Hypothesis, question 4 to his Behavioral Hypothesis, and question 5 to his Behavioral Intention. Support for the theoretical assumption that the Reinforcement Hypothesis and the Behavioral Hypothesis should affect performance only through the Behavioral Intention was found in an investigation by Dulany (1962).

The categorization of the Ss' reports to these questions further defines the nature of the self-instructional sets. These categories are largely self-explanatory but a comprehensive elaboration of their practical and theoretical relevance is discussed (Dulany, 1962). They are as follows:

Report of Behavioral Hypotheses: The subject . . .

I. names the correct response class and calls it correct or describes it as what he is supposed to do or as what _E_ wants him to do.

II. names as correct (or as what he is supposed to do or what _E_ wants him to do) some response class that is positively, but imperfectly, correlated with the correct response class.

III. does not name the correct response class as correct. He may report that he does not know the correct response, that he does not know whether there is a correct response, that there is not a correct response or he may report some uncorrelated and irrelevant response class as correct. (An irrelevant response class is one that is uncorrelated but not incompatible with the correct response class.)

IV. names as correct some response class that is uncorrelated but partially incompatible with the correct response class--e.g., right-left alternation or position preference when the correct response is randomly distributed right and left.

V. names as correct some incorrect response class that is negatively correlated with the correct response class.

Report of Reinforcement Hypotheses: The subject . . .

A. reports the significance of the contingent stimulus--that it signified that the preceding response was correct or what _E_ wanted or would agree with. It is described as having some selective reinforcement or information value, not as a general encouragement to continue.

B. reports the distribution but not the significance of the contingent stimulus--that it followed the response class _E_ designates as correct.

C. reports the occurrence but neither the significance nor the distribution of the contingent stimulus.

D. does not report occurrence of the contingent stimulus.

Report of Behavioral Intentions: The subject . . .

1. reports intention to produce the response class E designates as correct.
2. reports intention to produce some response class that is positively, but imperfectly, correlated with the correct response class.
3. reports no particular intention or reports intention to produce some irrelevant response class.
4. reports intention to produce some response class that is uncorrelated but partially incompatible with the correct response class.
5. reports intention to produce some response class that is negatively correlated with the correct response class (preceeding formulations from Dulany, 1962, pp. 113-115).

Dulany's formulations have been widely cited since they were first proposed. A classification by Locke (1966) of the four major types of questions used in verbal learning studies by different Es testifies to the usefulness of Dulany's theory. Locke's categories are (Type I) Purpose of the Experiment, (Type II) Response-Reinforcement Contingency, (Type III) Intention to get the Reinforcement or to Give the Correct Response, and (Type IV) Recall of Behavior. Throughout the subclassifications of these four main headings the link to Dulany's theory is apparent. Questions of Types II and III would be involved in the most direct test of the operation of the law of effect. Type II questions relate directly to S's awareness of the response-reinforcement contingency and Type III to his intentions. On the other hand, Type I questions are indirectly related to the law of effect. These questions are more cogent for inferences regarding awareness of the response-reward contingency. Locke states, "In effect, Type I questions are most directly pertinent to S's awareness of the demand characteristics of the experimental situation" (Locke, 1966, p. 295). Such questions ask whether S knows what E wants or expects him to do. Likewise, Type IV has little to do with the law of effect and is related only to memory of behavior.

The Relevance of a Recent Series of Conditioning of Meaning Experiments

A series of experiments related to the literature discussed thus far has been recently performed by Nunnally and associates (Nunnally, Stevens, & Hall, 1965; Nunnally, Duchnowski, & Parker, 1965; Parker & Nunnally, 1966; Nunnally, Knott, & Duchnowski, 1967; Knott, Nunnally, & Duchnowski, 1967; Faw & Nunnally, 1967; Hall, 1967; Nunnally, Duchnowski, & Knott, 1967; Nunnally & Faw, 1968). They have used conditioning devices such as a wheel spin or a slot machine which S sets in motion but over whose outcome S has no control. They have employed such dependent measures as verbal evaluation, selective attention, and reward expectancy as effects of schedule of reinforcement, delay of reinforcement, number of conditioning trials, delay of testing, massed versus distributed practice, age, and sex. All of these variables have been considered as they relate to the effect of rewards and punishments upon neutral objects.

In general, the results of these studies have yielded the following outcomes. Rewarded stimuli are preferred more than punished or neutral stimuli whether measured by adjectives assigned to the stimuli or by measures of selective attention (e.g., eye movements, pupillary response, and instrumental viewing responses). Such conditioning also leads to greater "expectation of reward" in other similar settings. This expectancy of reward is an increasing monotonic function of the percentage of reinforcement during conditioning. None of the effects were due to the age or sex of the children, nor to levels of massed or distributed practice.

Of particular interest were three additional findings: First, the effects were as strong for Ss tested five weeks later as for those tested the day following the end of conditioning sessions. Second, anticipation of reward had as strong an effect as that due to the actual pairing of neutral objects with rewards. In this situation Ss were told, "Tomorrow when the pointer stops on this make-believe word, you will win two pennies from this big pile. If the pointer stops on this make-believe word tomorrow, you will lose one penny and have to return it to the pile." Third, the effects were as strong for Ss receiving only five conditioning trials as for those receiving larger numbers of trials. In discussing this third finding, Nunnally, Knott, and Duchnowski (1967) attribute this "conditioning" effect to a gradual refinement of their procedures "to obtain the strongest and most rapid conditioning that could be accomplished" (Nunnally, Knott, & Duchnowski, 1967, p. 260). While the authors admit that verbal evaluation can be thought of mainly as a "cognitive" measure, they make no systematic attempt to assess Ss' hypotheses or intent and hold that "the conditioning apparently occurs in full force in early conditioning trials" (p. 261).

It appears reasonable, with a limited number of attributes, that five reinforcements may be enough for S to understand that the critical stimulus delivers pennies and adjust his preferences accordingly. This appears to be a more reasonable explanation for the phenomenon than the mechanism of "the strongest and most rapid conditioning that could be accomplished." That is, if cognitive processes are considered in such experiments the effects of conditioning may be slight. Such considerations are the primary concern of the present experiment.

Summary

Conditioning of meaning is a topic closely related to Thorndike's law of effect, which holds that the action of rewards can be automatic and independent of conscious perceptions. But Thorndike realized that persons are often aware of the response-reinforcement contingency, and he believed that such knowledge was beneficial to Ss' learning and performance.

Several studies were cited as an outgrowth of Thorndike's work. Although these experiments have clarified the issues and contributed to a better understanding of awareness and the automatic action of reinforcers, they have failed to resolve differences of opinion held by contemporary investigators. For example, Postman has opposed the Adams position which is supported by Dulany and Spielberger.

A series of recent studies was conducted to investigate verbal operant conditioning. From 1950 to 1958 little evidence was found to support the notion that Ss become aware of the reinforcement contingency. Since that time, however, the majority of investigators, who have employed intensive postinterview techniques, report little or no verbal conditioning can be attributed to Ss classified as unaware. A number of older studies such as that of Rees and Israel (1935) have been attacked on this basis.

Although several modified methods for assessing awareness have been proposed, the verbal report remains the technique most often used. It is pointed out that many children, old persons, brain damaged individuals, and mental defectives are unable to make such verbal reports although they show effects of conditioning.

Even classical-type semantic conditioning studies have had to deal with the awareness question. While some (e.g., Staats & Staats) report no evidence that Ss become aware of the CS-UCS contingency, others (e.g., Maltzman) indicate the reliable superiority of these Ss.

In such conditioning of meaning experiments it is probable that, through learning sets and general awareness, S forms perceptions about the experimental demands of the situation. These demand characteristics influence the mediation of the affective or connotative meanings involved.

Dulany has proposed a system for assessing Ss' hypotheses about the purposes, demands, and response-reinforcement contingencies. His system includes a theoretical framework for relating reinforcement hypotheses, behavioral hypotheses, and behavioral intent to performance, and an interrogation procedure yielding information concerning S's awareness of these mechanisms.

A recent series of conditioning of meaning experiments by Nunnally and his colleagues was described. They have found that rewarded stimuli are preferred more than punished or neutral stimuli. Verbal evaluation and selective attention have been viewed by those authors as being more aptly accounted for by the automatic action of reinforcement than by mediated awareness of the demands of the experimental situation.

CHAPTER III

PURPOSE OF THE EXPERIMENT

The present experiment is based upon the series of studies by Nunnally and his associates described earlier. The principal purpose of their investigations was to examine the effects of modified classical conditioning procedures on the development of the preferences of children. In their words, "The general purposes of this line of investigation are (a) to determine the treatment variables that influence the strength with which neutral objects become associated with rewards, and (b) to determine the forms of behavior toward the neutral object that are affected by the association formed between the neutral object and the reward" (Nunnally, Knott, & Duchnowski, 1967, p. 250).

An analysis of the procedure employed in the Nunnally, Stevens, and Hall (1965) Experiment I, suggested that the change (increase) in preference over trials might be attributed to cognitive factors, to conditioning, or to both. The present experiment was designed to separate these effects more clearly.

Nunnally et al. (1965) conclude that the "conditioning" procedure had marked effects on consequent verbal evaluation of the stimulus. Furthermore their _S_s, in postexperimental interviews, were found to be aware of reinforcement contingencies, but even with leading questions, a deliberate effort to "help the _E_" was not identified. Nevertheless, the authors admit that because of their Ss' low level of verbal facility and because of E's questionable interviewing procedures, "we presently do not know the forms of awareness and intention that accompany the experimental treatments" (Nunnally, Stevens, & Hall, 1965, p. 55).

Even if the most determined questioning failed to elicit verbal evidence of S's awareness of the experimental demands, this would still not be crucial evidence of lack of intention to help E. In addition, it has already been noted (p. 23) that E might encounter difficulties in employing such procedures because of children's limited verbal facility in expression. Furthermore, a verbal report is, itself, subject to modification by demand factors.

Accordingly, researchers are more convinced by evidence for or against cognitive control obtained via behavioral measures than they are by introspective report. Ideally, an experimental design with the discovery of awareness and of the effects of demand characteristics among its purposes should control for effects of repeated testing, which may be confounded with the manipulation of the number of reinforcements (one, two, or three conditioning sessions) as in the Nunnally, Stevens, and Hall (1965) study. Concomitantly, a design in which these variables are controlled would increase external validity (Campbell & Stanley, 1963). As Orne (1962) has pointed out, the extent to which the S's behavior is related to his awareness and intention, rather than the experimental variable, will in large measure determine both the extent to which the experiment can be replicated with modified demand characteristics and the extent to which generalizations can be drawn about the effect of the experimental variables in non-experimental contexts.

These considerations involve the question of whether the measurement of affect, in Nunnally's experiments, influences affective ratings by the subject. To be more specific, does an experiment in which four evaluations of conditioned stimuli are obtained provide Ss with cues (whether explicitly or implicitly) to the experimenter's hypotheses? Will the procedure in which four ratings are obtained yield different results from the procedure in which only one measure is taken? If the S's evaluation of rewarded stimuli does, indeed, differ as a consequence of manipulations in ratings, what mechanism can account for the difference? Rating procedures from one occasion to the next are similar in their structural characteristics. If a stimulus is rated with several adjectives reflecting similar evaluation, it might be expected that a corresponding degree of habit strength or probability of responding with similar adjectives in future ratings is increased by the mere frequency of contiguous association between attribute and adjective. Does learning-to-learn across repeated structurally similar ratings function to facilitate increased awareness of the experimental demands yielding concomitant inflation of the evaluation? The greater the number of rating sessions involved in an experiment, the greater is the opportunity for habit strength of previously chosen adjectives and learning-to-learn to operate. If such mechanisms are operating, what is the trend of the function (positively or negatively accelerated, decelerated, linear, or quadratic)?

Hypotheses

On the basis of an analysis of the Nunnally, Stevens, and Hall (1965) study in accordance with the above rationale, several additional assumptions can be made for the present experiment. Thus, during the course of a seven day conditioning of meaning experiment there are ostensibly two processes involved: rewards are (or are not) associated with specific (originally neutral) stimuli, and ratings of the stimuli are made on each rating day. The procedure has several possible consequences: (a) the conditioning trials may have no effect; (b) stimuli may acquire meanings through frequent contiguous association with the reward; (c) the S may learn, with increasing precision "what the experimenter wants him to do," and he may (or may not) act accordingly. These consequences are diagrammed in Table I.

Information that facilitates the formation of hypotheses or perceptions of experimental demands in a two-phase experiment can be obtained by S through the situation surrounding the administration of reinforcements and through the nature of the questions involved in the rating task. Thus, an analytical experiment is required to separate the effects of conditioning per se from the possible effects of sequential or multiple ratings of stimuli. One alternative is the design chosen for the present study in which manipulations of constant (the same stimuli on all conditioning and rating days) and variable stimuli (stimuli changed after each rating day) are crossed orthogonally with variable numbers of evaluations. In the rating manipulation the number of conditioning trials (reinforcements and

TABLE I

SOME POSSIBLE EFFECTS IN A LONG-TERM SEMANTIC CONDITIONING EXPERIMENT

Effect		Process
Semantic Conditioning	No Effect	1 Ineffective Conditioning
		2 Combination of + and - Factors
	Positive Evaluation	3 Automatic Action of Rewards
		4 Partial Understanding of the Response-Reinforcement Contingency
		5 Cognitive Perception of Experimental Demands
		6 Combination of 3, 4, and 5
	Negative Evaluation	7 Deliberate Negative Intention
		8 Misinterpretation of the Reinforcement
		9 Combination of 7 and 8

days) can be kept constant, but the number of ratings can be varied. In the stimulus manipulation the conditioned stimuli can be kept constant over all trials (days) or can be changed from trial to trial (from one conditioning session to the next).

Accordingly, if, in the manipulation of number of ratings, the rating on the last trial varies as a function of the number of evaluations made, the hypothesis is that ratings contribute to the learning of demand characteristics and increased awareness. Conversely, if the rating on the final day does not vary as a function of number of ratings, but the rating does vary as a function of trials, then the hypothesis is that number of reinforcements (reinforced trials) is a primary determinant of changes in meaning. (Note, however, that this still does not separate the effects of reinforcements themselves from the effects of ratings on demand characteristics, an interesting notion and one that is worthy of further investigation, but unfortunately is beyond the scope of the present study.)

The specific experimental hypotheses investigated in this experiment, and based on the above rationale, are as follows:

Hypothesis I. Preference can be conditioned by associating stimuli with secondary reinforcers.

Although this hypothesis has been supported many times in previous experiments, it provides a necessary baseline in the present study. Hypotheses II and III are central to the purposes of the present investigation.

The secondary reinforcers in this case are marbles and toys which are assumed to serve a function similar to trinkets in a study by Schroder (1956). In his experiment it was demonstrated that formerly neutral tokens acquired reinforcement value for elementary school children. The tokens were inserted into a vending machine which dispensed trinkets, and different schedules of reinforcement produced reliable effects in the conditioning of preference for the token. Similarly, in the present experiment, it is assumed that preference for a critical stimulus can be acquired through association with small toys and marbles.

Hypothesis II. The magnitude of the preference for a given stimulus is directly related to the number of ratings of that stimulus made by S, as well as a function of the number of reinforcements associated with the stimulus.

It appears reasonable that during an experiment in which several similar evaluations are obtained on the same stimulus, the *S* would acquire considerable sophistication about the demands of the experiment by the time it was concluded. This process would be analogous in some respects to learning-to-learn in which the process of making several ratings of the same (similar) stimulus parallels the repeated solution of the same (similar) problem. Thus, through nonspecific transfer of rating skills and other skills required for participation in a given experiment, the *S* becomes more efficient at solving that kind of problem. More specifically, over several problems *S*s acquire the rule that *E* "wants the reinforced one to be called 'good'" or some comparable learning set.

There is an alternative hypothesis which must also be considered here. Repeated ratings of the same stimuli over conditioning days may yield an increase in CS preference as it did in the Nunnally, Stevens, and Hall (1965) study. However, the group which rates the stimulus only once (on the final day) may evaluate it just as positively on that final day as does the group which makes multiple ratings. If these results are obtained, then it is reasoned that: *the changes in preferences are due primarily to reinforcements rather than ratings or experience with the same stimuli*.

Hypothesis III. *Ratings of conditioned stimuli in which all stimuli are varied from one conditioning session to the next will increase in polarization over the course of the experiment due, primarily, to the effects of rating*.

This, of course, is the important hypothesis regarding the effects of the rating procedure on the acquisition of knowledge about demands. Similar ratings of different stimuli are assumed to consist of repeated problems of the same type. Although a companion to Hypothesis II, Hypothesis III refers to the effects of repeated ratings on *different* stimuli, whereas Hypothesis II refers to the effects of repeated ratings on the same stimuli. Thus, to the extent that frequency of association between a stimulus and reinforcement has a strengthening effect on preference, independent of ratings, the function hypothesized for repeated ratings of different stimuli (Hypothesis III) should be weaker than that hypothesized for repeated ratings of the same stimulus (Hypothesis II).

Again, the alternative possibility is that no change will occur in evaluations of the stimuli which are changed on each new conditioning day (i.e., ratings are the same from day to day or from set to set) but are higher than with experience (without reinforcement) only. If so, it can be hypothesized that changes in preference are due primarily to reinforcements rather than to ratings or experience with different stimuli.

The conditions related to manipulations of stimuli thus, can be seen to control for learning-how-to-learn and warm-up-effects, and the relative contribution of ratings and reinforcement to the evaluation of stimulus meanings can be determined.

A postexperimental procedure may be used to investigate the identification of Ss' behavioral and/or reinforcement hypotheses as they are correlated with experimental variations. Thus, S's awareness of the experimental demands should become apparent if he is required to change places with E, and to administer reinforcement to the E who now plays the role of S. Ss who have correct hypotheses about the nature of the experiment should be superior to those whose perceptions are incorrect in performing the tasks associated with E's role. Conversely, Ss who have formed incorrect reinforcement hypotheses will fail to perform accurately the reinforcing role when the correct stimulus appears, and Ss with incorrect behavioral hypotheses will reward inappropriate stimuli. Although these several interpretations of the consequences of the role-playing procedure may be difficult to separate statistically, the procedure holds the promise of being a useful alternative to verbal interview for young children. Only reinforcement and behavioral hypotheses are of concern in this procedure since, with young children, who are usually quite cooperative in an experimental setting, it is assumed that Ss select their answers to agree with their behavioral hypotheses; i.e., behavioral hypothesis and behavioral intent are consonant. In addition a modified Dulaney (1962) postexperimental interview should provide corroborative and clarifying awareness data to complement those of the role-playing procedure.

Summary

Analysis of Experiment I performed by Nunnally, Stevens, and Hall (1965) suggested that increased preference for a critical stimulus over trials might be attributed to cognitive factors, to conditioning (reinforcements), or to both. The present experiment was designed to investigate the independent effects of these two variables on the development of preferences in conditioning experiments.

Accordingly, the following primary hypotheses were developed: Preferences can be conditioned by associating stimuli with secondary reinforcers. The magnitude of the preference for a given conditioned stimulus is directly related to the number of ratings of that stimulus made by S, as well as a function of the number of reinforcements associated with the stimulus. Ratings of conditioned stimuli in which all stimuli are varied from one conditioning session to the next will increase in polarization over the course of the experiment due, primarly, to the effects of rating. However, the possibility that the Nunnally et al. (1965) results might be due solely to reinforcements rather than ratings or experience was retained as an alternative hypothesis.

Postexperimental measures of awareness were viewed as complementary to behavioral measures of cognitive performance inherent in the design.

CHAPTER IV

METHOD

The overall design employed in this investigation consisted of four levels of rating schedules orthogonally crossed with two levels of stimulus conditions. An additional control group, apart from the main design, was provided for the purpose of providing experience with the stimuli in the absence of reinforcement. In one variation of the stimulus conditions the stimuli that were associated with reinforcements were always the same from one training session to the next (Condition SS) throughout the experiment, and in the other, the stimuli differed among sessions (Condition DS). Each of the SS and DS conditions were orthogonally crossed with four rating schedules varied according to the number of rating days. Of the seven days in which the experiment was conducted, Group R4 rated stimuli on four days (days 1, 3, 5, 7); Group R3 rated stimuli on three days (days 3, 5, 7); Group R2 rated stimuli on two days (days 5, 7); and Group R1 rated stimuli on one day (day 7). *Ss* in Groups R3, R2, and R1 were administered filler tasks on days in which ratings were not obtained.

The total design was comprised of eight groups with a separate control group. The latter was a control for reinforcement, or experience-only group (Condition EO). *Ss* in this group viewed the same stimuli throughout the experiment, as did *Ss* in the SS groups, rated stimuli on all four evaluation days, and performed all aspects of the task on training session days, but were never provided reinforcements. The overall experiment covered seven consecutive school days (rating or filler activity on days 1, 3, 5, 7, and conditioning trials on days 2, 4, 6) for each S. A graphic representation of the design is displayed in Table II.

Subjects

The Ss were 108 first grade children (12 in each of nine conditions), 58 boys and 50 girls, from the Bellefonte Elementary School in Centre County, Pennsylvania. In addition, 20 first graders, 10 boys and 10 girls, from the Pleasant Gap Elementary School in the same county participated in a pilot study and in preliminary evaluation of the adjectives to be used in the experiment.

The 108 Ss were members of four classes and were assigned by reference to a table of random digits to the nine experimental conditions with the restriction that all classes be represented approximately equally in each condition. Following assignment of the first nine Ss to the nine conditions, the randomization procedure was begun again, i.e., randomization of Ss was recycled at N + 1 treatments. This distribution of classes among conditions is reported in the Appendix, Table XII. There was no evidence that Ss in the four classes differed systematically from each other with regard to the experimental task.

TABLE II

REPRESENTATION OF TREATMENT CONDITIONS

	Verbal Rating Days				
	1,3,5,7	3,5,7	5,7	7	1,3,5,7
Same Stimulus (SS)	Group SS-R4*	Group SS-R3	Group SS-R2	Group SS-R1	Group SS-EO
Different Stimulus (DS)	Group DS-R4	Group DS-R3	Group DS-R2	Group DS-R1	n = 12 N = 108

Groups	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7
SS-R4*	Rate	Cond.	Rate	Cond.	Rate	Cond.	Rate
SS-R3	Fill	Cond.	Rate	Cond.	Rate	Cond.	Rate
SS-R2	Fill	Cond.	Fill	Cond.	Rate	Cond.	Rate
SS-R1	Fill	Cond.	Fill	Cond.	Fill	Cond.	Rate
DS-R4	Rate	Cond.	Rate	Cond.	Rate	Cond.	Rate
DS-R3	Fill	Cond.	Rate	Cond.	Rate	Cond.	Rate
DS-R2	Fill	Cond.	Fill	Cond.	Rate	Cond.	Rate
DS-R1	Fill	Cond.	Fill	Cond.	Fill	Cond.	Rate
SS-EO	Rate	Expo.	Rate	Expo.	Rate	Expo.	Rate

Note.--The following abbreviations were used in the table above: Rate = verbal rating of the appropriate stimuli; Cond. = conditioning trials; Fill = neutral filler activity; Expo. = exposure without reinforcement; SS = same stimuli throughout conditioning and rating days; DS = different stimuli on conditioning days, i.e., all six stimuli change on days 2, 4, and 6; EO = experience only (no reinforcement).

*Replication of Nunnally, Stevens, and Hall (1965).

The ages of the children ranged from 72 to 93 months with a mean age of 81.05 months (see Appendix, Table XIII).

Apparatus, Rating Materials, and Filler Materials

Different materials were used in each of the three phases of the experiment: the training apparatus was employed for the conditioning sessions; stimulus blocks were used for the rating task; and puzzles provided filler tasks for unscheduled rating days.

Conditioning apparatus. The training apparatus (see Figure 4) consisted of a small, four-legged, circular-top table with a wheel-spin device.[1]

The table top of this apparatus was a circular board 28 inches in diameter which stood 16 inches from the floor and was covered with green felt material. A circle of 18 upright pins was located four inches from the outer edge. The stimuli (three drawings of one set, or 18 stimuli) were drawn in black ink on white cardboard rings in which holes were punched to allow them to fit over the pins and lay flat on the table. These rings had outside diameters of 28 inches and inside diameters of 19 inches. Each of the two cardboard rings had one set of stimuli on the front and one set on the back thereby allowing E to change stimuli between Ss rapidly.

---

[1] The primary basis for the decision to use the spin-wheel apparatus rather than the slot-machine was a personal communication from Nunnally that indicated the spin-wheel game to be a versatile piece of apparatus for studies of conditioning reward value.

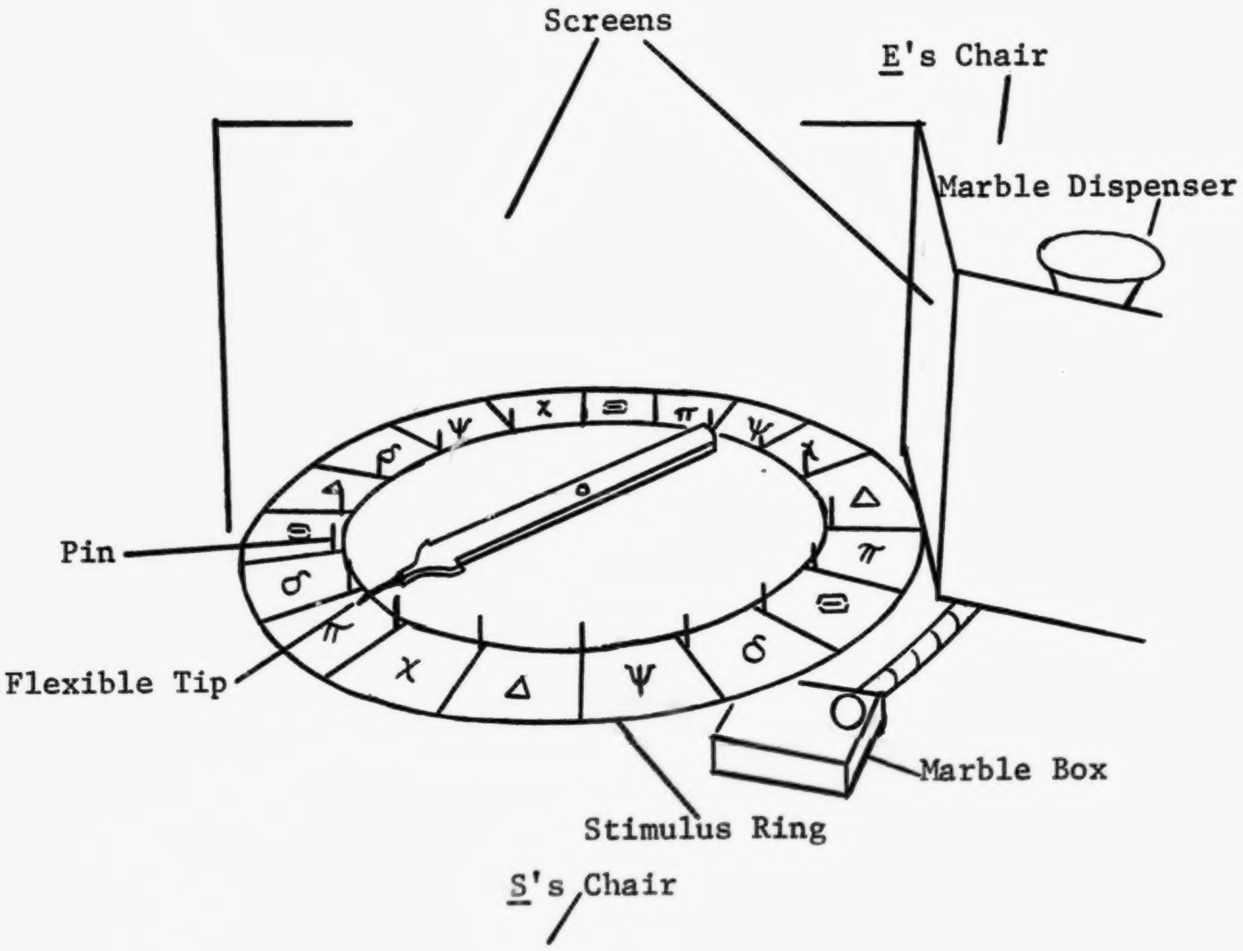

Screens
E's Chair
Marble Dispenser
Pin
Flexible Tip
Marble Box
Stimulus Ring
S's Chair

Figure 4. The surface of the conditioning table.

The center of an 18" x 2" x 1" formica arrow with a 1-1/8 inch flexible, plastic pointer attached to the arrow's head was mounted on a pivot in the center of the table. The pivot consisted of a 9" x 3/8" threaded steel shaft which was firmly secured to the center of the arrow, ran through a ball-bearing mounted in the table center, and extended 7-1/2 inches into a hollow wooden column. This column extended downward from the table top to within 3-1/2 inches of the floor (a length of 12-1/2 inches with a diameter of 3 inches). It was attached to the legs of the table by wooden spokes. The end of the shaft was seated in a bearing at the bottom of the column. Thus, when the arrow was spun, the pointer of the arrow brushed the pins thereby serving as the primary friction to retard the spin, since the entire threaded shaft turned easily.

A friction device by which E could control the stimulus on which the arrow stopped was also built into the apparatus. This device consisted of a 6-1/2" x 1/4" wooden dowel extending horizontally through holes in a leg and the column to engage the threaded shaft. A cable and foot pedal allowed E to apply varying amounts of pressure to this center shaft, thereby indetectably slowing the arrow until it stopped at the desired position. This device was used only when it was apparent that S was either going to exceed or fall short of the prescribed time for a conditioning session (see Procedure section).

A black, open-top box measuring 4-1/2" x 4-1/2" x 2" was mounted to the edge of the table top at S's right hand to collect the marble reinforcements for rewarded stimuli. A Gerbrands

Automatic Marble Dispenser was mounted on a 10" x 12" platform attached to the right side of the table (as viewed by S). The marble dispenser was operated by remote control by E, who viewed the Ss' spins throughout the trials and dispensed a marble each time a spin stopped on a stimulus to be rewarded.

The Ss sat in a child's chair at the table, and its size allowed them to comfortably perform whatever tasks were required. A system of screens blocked from S's view, the E (on conditioning days), the marble dispenser, and the friction control device. Routine maintenance such as replacing flexible tips and oiling the conditioning apparatus was performed periodically during the duration of the experiment.

Rating materials. The rating materials consisted of a cover for the conditioning table and stimulus blocks. The cover concealed the pins, arrow, and ring, thereby providing a flat surface on which the rating blocks or filler apparatus could be manipulated. The 28 rating stimuli were drawn in black ink on white 2" x 3" cards glued to black 2" x 3" x 2" wooden blocks (see Table III).

Filler materials. A jigsaw puzzle replaced the rating procedure for Ss who were not scheduled to rate the stimuli on a particular day. These puzzles were recommended by the manufacturer as being appropriate for children of age six. They were brightly colored picture puzzles of heavy fibreboard with loosely fitting pieces.

TABLE III

FOUR SETS OF STIMULI FOR CONDITIONING AND RATING

Set	Regular Stimuli	Extra Stimulus
A	Δ χ π δ Ξ Ψ	Λ
B	φ σ Ω ι β λ	ξ
C	ω Γ Z γ κ Θ	H
D	α μ ο Π τ Σ	η

Stimulus Materials

The stimulus materials for the conditioning and rating tasks were 28 (four sets of seven) upper and lower case Greek letters.[2] Shepherd (1957) indicated that such stimuli are discriminable by first grade children, and although the stimuli almost certainly have some association value, even for young children, they are relatively neutral in affective value before conditioning trials.

Each set of stimuli consisted of six standard stimuli and one "extra" stimulus which did not appear on the conditioning ring but was used in rating trials. On the four conditioning rings each of the six standard stimuli was drawn at three random positions about the ring with the restrictions that one complete order of the six stimuli must appear before the next began and that the same stimulus could not be adjacent to itself. Thus, there were 18 possible stopping places for the arrow. The stimuli were drawn over a 1 inch high template with 1/8 inch black ink lines. All stimuli faced S as he viewed the conditioning apparatus. The 28 rating blocks were stimuli identical to the 24 standard stimuli of the conditioning rings plus the one "extra" stimulus for each set, A, B, C, D. The four sets of stimuli (A, B, C, D) are listed in Table III.

---

[2] It was decided not to use geometric figures for stimuli, as did Nunnally, Stevens, and Hall (1965), since 28 orthogonally related figures could not be devised without resorting to random shapes.

Measures

Two measures of verbal evaluation of the stimuli were taken. Adjective assignment ratings similar to those used by Nunnally, Stevens, and Hall (1965) followed by a forced-choice adjective rating of the stimuli were given by S on each scheduled rating day. To supplement information about perceived demands and awareness which could be inferred from the design per se, two postexperimental measures of awareness were also administered.

Adjective rating. The adjective rating task required S to assign 14 adjectives to six stimuli or "pictures." The six rating blocks for this measure included the pay-off stimulus and four non-rewarded stimuli from the standard set as well as the "extra" stimulus for that set. This extra stimulus, which did not appear among the conditioning stimuli, was inserted in the rating in place of a randomly selected, deleted stimulus in order to determine the pattern of preference for such a stimulus. The 14 adjectives, (see Appendix, Table XIV) read in random order, to be assigned to the blocks were those used by Nunnally et al. (1965) with the exception that "mean" was substituted for "poor."[3]

---

[3] The basis for this substitution emerged in the pilot work for the experiment. In a preexperimental evaluation of 20 adjectives all of the Nunnally et al. adjectives received an evaluation which was highly polarized in the proper direction except "poor." A mean of 3.30 for "poor" as opposed to 4.55 for "mean" was registered on a five step semantic differential evaluation scale by 20 first grade pupils. Four groups of five Ss each marked an X in one of five boxes labeled "very good, a little good," blank, "a little bad, very bad." Ss received only one series of boxes on a slip of paper at a time, rated one word, and returned that slip to E before receiving another. Instructions and procedures for this

Verbal forced-choice rating. A verbal evaluation based on a forced-choice rating of the same stimuli involved in the adjective rating was also obtained. In this procedure, *S* applied either the positive or the negative adjective from each of seven pairs of bi-polar evaluative adjectives, presented orally by *E*, according to his evaluation of the stimulus before him. The bi-polar pairs (see Appendix, Table XIV), were constructed from the 14 adjectives used in the adjective rating task. They were read to *S* to be applied to each stimulus used on a given rating day.[4]

Postexperimental awareness measures. The first postexperimental measure was concerned with obtaining behavioral evidence of *S*'s reinforcement and behavioral hypotheses. It involved a reversal of roles in which *E* sat in *S*'s chair and spun the training arrow. *S*'s task was to play the role of *E* in the reinforcement situation. A correct reinforcement hypothesis was defined operationally as *S* ejecting marbles when the arrow stopped. A correct behavioral hypothesis consisted of ejecting a marble each time the arrow stopped at the CS. Additional criteria are discussed later.

---

preexperimental evaluation were adapted from Helper (1965). A replica of the semantic scale appears in the Appendix, Table XV, and the means are reported in that same table for this pilot evaluation.

[4] It was found that this measure was not sensitive enough to differentiate treatments. *Ss* in all conditions tended to make extreme ratings involving all positive or all negative adjectives. For example, a number of *Ss* rated all stimuli in the experiment totally positive. Since this measure did not yield fruitful results it will not be considered further in subsequent sections of this report. The results were as follows: $\chi^2(6) = 8.06$, $p > .05$ for a 3 x 4 matrix (SS, DS, and EO by 0 or 1, 2 or 3, 4 or 5, and 6 or 7 positive adjectives assigned by forced choice to the CS on the final rating day), and $F(8, 99) = 0.93$, $p > .05$ for positive minus negative adjectives assigned on the final day.

The postexperimental awareness interview also focused on the measure of S's reinforcement and behavioral hypotheses. The interview was adapted from Dulany's (1962) theoretical framework. Replies to the following questions and probes were recorded:

1. a. Did you notice whether anything happened during the time you were spinning the arrow? b. (What?) c. When the arrow would stop what happened?

2. a. Did you come to think it was just anytime or did it follow something in particular? b. (What?)

3. a. Did you come to think there was or wasn't any purpose or reason for having the marble in this game? b. (What?) c. Did the marble come out when the arrow pointed to one particular design or picture? d. Which one?

4. a. Did you come to think that there was anything you were supposed to say, or point to with the blocks in order to be right--something I wanted you to say or point to? b. (What?) c. Did you think there was or wasn't any kind of right answer?

5. a. How did you go about deciding which block to point to? b. If I said "Good" which one would you point to. . . . Why? c. Did you try to give the good words to the design or picture that got the marble or did you think about that? d. When did you push the button to give me marbles today?

Procedure

Each S was individually escorted to and from the experimental room in the school by E on each of the seven consecutive school days in which he participated in the experiment. The room was an attractive speech therapy room which was no longer in use for that purpose. S was seated at the apparatus table for each session. After a few introductory remarks the appropriate session was begun.

Before the first session of the experiment each *S* was shown the randomly assigned (by random digits) set(s) of stimulus blocks which he would eventually rate. In order to reduce variability due to inherent preference or dislike for a particular Greek symbol, *S* was instructed to point to the one that he liked "most" and the one which he liked "least" in each set as *S*s did in the Di Vesta (1962) study. Prior to this procedure, one stimulus from each set of standard stimuli had already been assigned at random as a "pay-off" stimulus to be rewarded during conditioning of a given *S*. If this assigned pay-off stimulus was either the "most" or "least" preferred figure, then a random assignment from the remainder of the stimuli was made.

On the first experimental rating day (e.g., day 1 for Groups SS-R4 and DS-R4, but day 3 for Group DS-R2; see Table II) the six appropriate rating blocks were laid on the table from left to right in front of the child.

S was instructed as follows: "Today we are going to play a word game (or, the word game again). I will call out a word, and you tell me which picture or design the word fits. You do not have to take turns with the blocks. Each time I say a word you point to the picture that you think fits that word the best." *E* then said, "The first word is ______. Which one do you think fits _____?" If *S* did not understand the word "fits," then "goes with" or "is" was used in its place. After *S* had assigned two adjectives to the blocks, the positions of the blocks were shuffled to prevent any bias which might occur on the basis of the positions of blocks on the table. The blocks were rearranged after every other adjective was assigned while S looked away from the blocks at a black paper circle attached to the wall at his left. E sat to the right of and slightly behind S for the rating phase of the experiment.

Immediately following the adjective assignment part of each rating session, the forced-choice semantic rating was administered by exposing one block at a time and instructing S, "Now we want to use just one block. I'll say two words, and you tell me which one fits this design or picture." E then inquired, "Is this picture pretty or ugly?" After S replied, E then asked, "Is this picture worst or best?" The procedure was continued with choices to be made from these pairs: good--bad; nice--awful; beautiful--messy; and happy--mean. Further, the procedure was repeated for each of the six blocks. The blocks were exposed in random order, and the order in which E said the pairs of adjectives (e.g., good--bad or bad--good) was varied throughout.

In order to control for the amount of time spent in the experimental setting, Ss who were not scheduled to make a rating on a particular day engaged in a filler task. They were seated at the rating table and told, "Today I want to see how many pieces of this puzzle you can put together. Do you like to do puzzles?" The answer was invariably "yes." E then said, "Good! Go ahead and try this one." If Ss finished assembling the jigsaw puzzle before the time (approximately ten minutes) necessary to rate the stimuli had expired, E engaged the S in questions and conversation about the picture he had just completed until the allotted time had elapsed.

On the first conditioning day S was seated at the conditioning table and instructed as follows: "Now, here is a different game you can play today. Give the arrow a spin and see what happens. Each time the arrow stops give it another spin." Approximately 0.5 second after the flexible pointer came to rest between the pins for the rewarded stimulus assigned to that S, a marble was dispensed by E. When S had received his first marble he was asked, "Do you see that marble? . . . Come over here and see these toys. [An assortment of 15 to 20 inexpensive toys such as balls, jacks, paint boxes, cars, horses, and rings was displayed on a library type table.] You will have a chance to try to earn one of these presents to keep." He was shown a card with six scholastic stars pasted on it, told that it took that many stars on his paper to "earn a present," and that he could earn two stars that day if he got 20 marbles in the box. These rules were reviewed for S on each conditioning day (2, 4, and 6). Both SS and DS children then continued to spin the arrow without further interruption until 20 reinforcements had been received, a figure based on a pilot study and on the Nunnally et al. (1965) study which suggested that a schedule of 20 reinforcements in 20 minutes should be effective. The preselected CS was always reinforced, and no other stimuli were ever rewarded. Ss were permitted to count their marbles as often as they wished. The length of the conditioning session was approximately 20 minutes, and the friction control device was employed by E only in cases where this approximate schedule of one reinforcement per minute was not being maintained by S's chance spins.

S helped E paste his two stars to the back of his data sheet at the conclusion of each conditioning session. All Ss received conditioning trials and twenty reinforcements on the second, fourth, and sixth days with the exception of Group SS-EO. In the control condition (EO) the arrow stopped twenty times on the selected stimulus, but no reinforcements were given. The instructions for this group were modified accordingly to indicate that S could earn his two stars by continuing to spin the arrow until E told him to stop.

After the final rating on the seventh day (for all Ss except those in the control for reinforcement Group SS-EO) E showed and demonstrated to S the remote control button which dispensed the marbles. S was then given instructions for the role reversal phase of the experiment. "Now I want to trade jobs with you. I want to sit there in your chair and play like I am you and spin the arrow. I want you to sit here in my chair and play like you are me and do exactly what you think I did while you were spinning. We want it to be just like it was yesterday when you were spinning the arrow." E then sat in Ss chair and spun the arrow a minimum of 10 times. If the pointer had not stopped upon the rewarded stimulus twice in 10 trials, E continued to spin until it had. At this juncture if S had correctly reinforced both occurrences and no other stimulus, or failed to reinforce both occurrences, or had reinforced every spin then this role reversal phase of the experiment was terminated. If, however, S reinforced only one of the occurrences of the critical stimulus and several other stimuli, the spins were continued until

S's reinforcement hypothesis and intention were apparent. After the completion of this task, S returned to his chair, and the post-experimental interview was conducted.

Following this interview, S was given the toy of his choice with a letter of information and thanks to the parents (see Appendix, Figure 6). The toy chosen was replaced on the table from E's stock before the next S arrived. Although, naturally some Ss were absent on certain experimental days, they were continued at the proper place when they returned. The greatest number of extra days missed by an S was five (see Appendix, Table XVI for means).

CHAPTER V

RESULTS

A preference score for each stimulus was obtained on each rating day by subtracting the number of negative adjectives from the number of positive adjectives assigned to it. Thus, a preference score for the critical (reinforced) stimulus was obtained for each S on days when adjectives were assigned to the stimuli.

The initial analysis performed was a one-way analysis of variance of the preference scores for all groups obtained on the final, or seventh, day of the experiment. This analysis was made to obtain an overall estimate of the error variance based on all groups in the experiment, as well as to determine whether any significant treatment effects were obtained. The analysis is summarized in Table IV where it can be seen that the effect due to Treatments is significant ($F = 2.77$, $df = 8/99$, $p < .01$). The means and standard deviations of the preference scores employed in this analysis are presented in Table V, and are graphically displayed by the final point of the curves in Figure 5a and b. The reader will also note in Figure 5a that the present study yielded an exact replication of the results found in the Nunnally, Stevens, and Hall (1965) study where the comparable conditions were employed. An $F_{max}(8, 11)$ test for homogeneity of variance of the eight experimental groups yielded a ratio of 3.27 which was not significant ($p > .05$). When the control

TABLE IV

ANALYSIS OF VARIANCE OF FINAL PREFERENCE SCORES

Source	df	MS	F	p
Treatment	8	10.63	2.77	<.01
Error	99	3.84		
Total	107			

TABLE V

MEANS AND STANDARD DEVIATIONS OF FINAL PREFERENCE SCORES

Stimulus Condition		Rating Condition				All Groups	Control (EO)
		R4	R3	R2	R1		
SS	$\overline{X}$	2.92	2.17	1.25	1.08	1.85	0.08
	SD	1.93	1.91	1.66	2.35	1.96	0.90
DS	$\overline{X}$	0.00	0.92	0.50	0.50	0.48	
	SD	1.95	2.94	1.62	1.73	2.06	
All Groups	$\overline{X}$	1.46	1.55	0.88	0.79	1.17	
	SD	1.94	2.43	1.64	2.04	2.01	

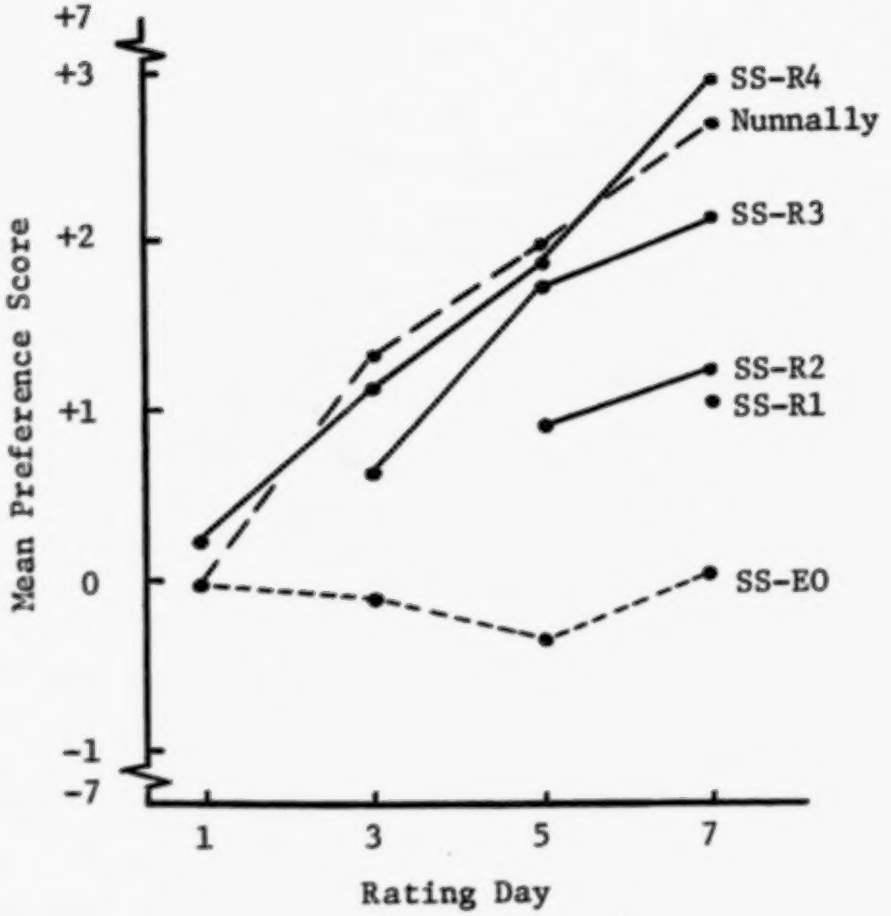


Figure 5a. Mean preference scores for SS groups and for the Nunnally, Stevens, and Hall (1965) group over all rating days. The latter group is comparable to the SS-R4 group in the present study.

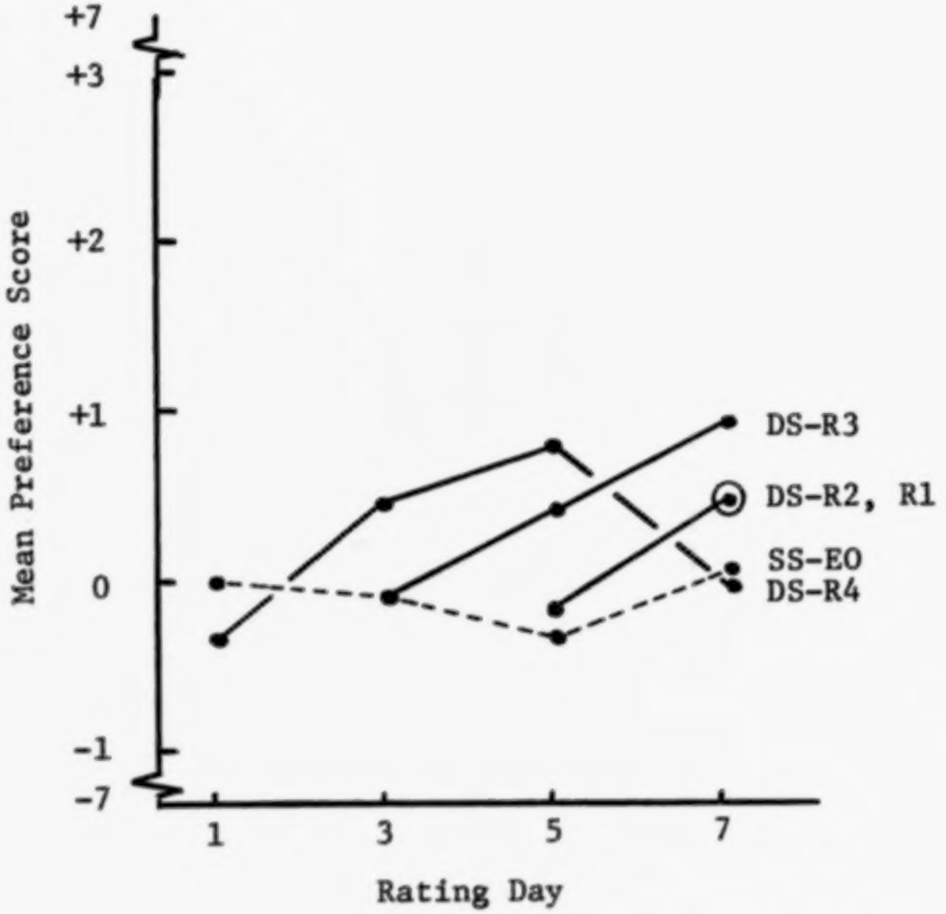


Figure 5b. Mean preference scores for the DS groups and for the Control (SS-EO) group.

condition EO was included, an $\underline{F}_{max}(9, 11) = 10.65$, $\underline{p} <.05$ was obtained. However, this was not considered to be sufficiently high to warrant transformation of the data (see discussions by Box, 1954 and Winer, 1962).

The significant overall effect of Treatments resulted primarily from a tendency for $\underline{S}$s in Group SS-R4 to evaluate the CS more positively than $\underline{S}$s in Group SS-EO or in any of the DS groups. _A posteriori_ multiple-comparison tests by the Tukey "a" procedure (Winer, 1962, p. 87) yielded significant differences ($\underline{p} <.01$) between final ratings by Group SS-R4 and those for Group DS-R4 as well as Group SS-EO. Group SS-R4 was also found to differ significantly ($\underline{p} <.05$) from groups DS-R1 and DS-R2 on the final ratings. The less conservative Newman-Keuls (Winer, 1962, p. 80) procedure yielded an additional significant ($\underline{p} <.05$) difference between Group SS-R4 and Group DS-R3. The results of the multiple-comparison tests are summarized in Table VI.

The second analysis of the preference scores on the final rating day consisted of a 2 x 4 factorial analysis of variance involving the two levels of stimulus variation (SS and DS) and the four rating schedules (R4, R3, R2, R1). This analysis is summarized in Table VII. As would be expected on the basis of the previous analyses, including the multiple-comparison tests, a significant effect due to Stimulus Variation, $\underline{F}\ (1/91) = 9.91$, $\underline{p} <.01$, was obtained. However, neither the effects due to Rating nor to the Stimulus Variation x Rating Schedule interaction were significant ($\underline{p} >.05$).

TABLE VI

MULTIPLE COMPARISONS OF FINAL PREFERENCE SCORES BY ALL GROUPS

Comparisons	*p*	Test
SS-R4 and DS-R4	<.01	Tukey
SS-R4 and DS-R3	<.05	Newman-Keuls
SS-R4 and DS-R2	<.05	Tukey
SS-R4 and DS-R1	<.05	Tukey
SS-R4 and SS-EO	<.01	Tukey
All other comparisons	>.05	Tukey

TABLE VII

ANALYSIS OF VARIANCE OF FINAL PREFERENCE SCORES FOR STIMULUS AND RATING SCHEDULE CONDITIONS

Source	*df*	MS	*F*	*p*
Stimulus Condition (SS and DS)	1	42.67	10.11	<.01
Rating Schedule	3	3.15	0.75	>.05
Interaction (Stimulus x Rating)	3	6.94	1.64	>.05
Error	88	4.22		
Total	95			

An examination of the means of the conditions plotted in Figure 5 reveals that the preference ratings for the critical stimulus of *S*s in all SS groups is superior to the preference ratings of even the highest DS group. That is, although some of the apparent differences are not significant statistically, on any given rating day no mean preference rating for any of the experimental SS groups is equal to or less than the highest mean preference rating by any of the DS groups. If the curve for Group DS-R4 were extrapolated, it would approximately coincide with that for Group SS-R1; however, the sharp drop from day 5 to 7 was due entirely to the ratings of 3 of the 12 *S*s in Group DS-R4. An attempt to identify the reason for this change was unsuccessful.

The third analysis was related to the assumption that SS-R4 and DS-R4 groups should increase in preference for the critical stimulus over four trials (ratings). It will be recalled that Group SS-R4 received 60 reinforcements for one CS, Group DS-R4 received 20 reinforcements for each of three CS's and a given *S* in Group SS-R0 received no reinforcements for his assigned stimulus. Thus, the analysis sought to clarify the effects of repeated ratings and number of reinforcements on preference for the critical stimulus.

The data were analysed by an analysis of variance with three treatments (Groups SS-R4, Group DS-R4, and Group SS-R0) for the between variable and the four ratings as the within variable. The effect due to Treatments was significant, F (2/33) = 7.30, $p < .01$. In addition, the effects due to Rating Trials ($F = 3.08$, $df = 3/99$, $p < .05$) and to the Treatment x Rating Trials interaction ($F = 2.32$, df = 6/99, $p < .05$) were significant. This analysis is summarized in Table VIII and means for all ratings appear in Table IX. A trend analysis of the Group SS-R4 curve yielded a significant ($p < .01$) linear trend as was found in the Nunnally, Stevens, and Hall (1965) Experiment I. Quadratic and cubic tests were not significant ($p > .05$).

The above analyses show significant effects due to stimulus variation, thereby clearly supporting the hypothesis that frequency of association between a neutral stimulus and a reinforcing stimulus demonstrably influences preference for the former. In addition, repeated evaluation of all stimuli in the SS-EO control condition was consistent across rating days for the majority of Ss in that group. Computation of Kendall's Coefficient of Concordance for each S across all stimuli revealed significantly reliable ratings for seven of the 12 Ss ($p < .01$ for five Ss and $p < .05$ for two Ss). The Coefficients of Concordance ranged from .05 to .86 with a mean of .49.

However, the analyses described up to this point provide little support for the hypothesis that the number of ratings tends to increase S's sensitivity to experimenter demands and consequently should enhance preference for the stimulus over and beyond that effect due to association with reinforcement. Despite the lack of significance for this effect via analysis of variance procedures it is readily apparent that the order of the obtained preference means corresponds exactly with the predicted order based on ratings for all of the SS groups and for all but one of the DS groups. Furthermore, the reader's attention is called to the argument that the analysis

TABLE VIII

MIXED ANALYSIS OF VARIANCE OF PREFERENCE SCORES BY GROUPS SS-R4, DS-R4, and SS-EO ACROSS FOUR RATING TRIALS

	df	MS	F	p
Between Subjects				
Treatment	2	36.34	7.30	<.01
Error	33	4.98		
Within Subjects				
Repeated Ratings	3	7.19	3.08	<.05
Interaction (Treatment x Rating)	6	5.89	2.52	<.05
Error	99	2.34		

TABLE IX

MEAN PREFERENCE SCORES FOR ALL GROUPS ON ALL RATING DAYS

Group	Rating Day			
	1	3	5	7
Group SS-R4	0.250	1.167	1.917	2.917
Group SS-R3		0.667	1.750	2.167
Group SS-R2			0.917	1.250
Group SS-R1				1.083
Group DS-R4	-0.333	0.500	0.833	0.000
Group DS-R3		-0.083	0.417	0.917
Group DS-R2			-0.167	0.500
Group DS-R1				0.500
Group SS-EO	0.000	-0.083	-0.333	0.083

Note.--Blank spaces were days in which ratings were replaced by filler activity for the groups concerned.

of variance statistic is not particularly appropriate for testing the ordered alternative hypothesis because of its insensitivity to the specific hypothesis it is desired to examine here.

Thus, as Jonckheere (1954) indicates:

> The customary one-way analysis of variance does not satisfy this demand: the F-ratio is independent of the order in which the group means occur. Furthermore, attempts to combine the probabilities yielded by an F-ratio and some coefficient of rank correlation between the expected and the obtained order of the group means are faced with a number of difficulties. . . . A further aspect of many experiments is that there appears to be no suitable metric by which the different treatments to be investigated may be characterized. All that an experimenter is able to assert particularly in psychology, is that they may be ranked in some order. . . . We have then the problem of assessing whether the results obtained from a one-way analysis of variance design are "significant" and in conformity with a hypothesis as to the rank order for the size of the effects yielded by different treatments.
> (Jonckheere, 1954, pp. 133, 134)

There can be no doubt that according to the theoretical framework for the present study that preference scores should correspond directly with (fall in the same order as) the number of ratings made. Accordingly, the Jonckheere (1954) distribution-free k-sample test against ordered alternatives was considered to be a completely justifiable and acceptable procedure for testing the null hypothesis.

The Jonckheere test for SS groups yielded an S of 410 which, when converted, resulted in a t (151) = 2.72, p <.01, when data from Group SS-EO were included. However, there may be reason for excluding the SS-EO group since, while it is a control group, it provides a control for reinforcement rather than number of ratings. For this reason another analysis was performed with the SS-EO group excluded and using only the ratings by Ss in SS Groups R1, R2, R3, and R4. The analysis of these data, appropriately ordered, yielded an S of 182, which converted to $\underline{t}$ (113) = 1.68, with $\underline{p} < .05$ for Q = 1. A similar analysis of the preference scores for the corresponding DS groups was not significant ($\underline{p} > .05$). Thus, an increase in the number of rating sessions contributes significantly to the ordering of the magnitude of preference scores in the SS groups but not in the DS groups. The computation of the S statistic and its $\underline{t}$ conversion are illustrated in Appendix, Figure 7.

The evaluation of the extra-experimental stimulus (see p. 67), which was inserted on rating days in place of a standard stimulus, deleted at random, was compared with the evaluation of a randomly selected standard stimulus for each $\underline{S}$. This comparison was made over all groups for the final rating session (seventh day). As can be seen from the means in Table X, no systematic pattern of preference or dislike emerged for the extra-experimental stimulus inserted on rating days or for randomly selected standard stimuli. The predominately negative values in that table are to be expected since a disproportionate number of positive adjectives were assigned to the critical stimulus. These data provide further support for the effects of the experimental manipulations on the development of preferences.

Postexperimental Awareness Data

The data for the role-taking situation in which $\underline{S}$ reinforced $\underline{E}$ were scored according to the following criteria: (a) A score of 2 was assigned if $\underline{S}$ reinforced both occurrences of the CS and no

TABLE X

COMPARISON OF FINAL MEAN PREFERENCE SCORES FOR THE EXTRA-EXPERIMENTAL STIMULUS AND THE STANDARD STIMULUS

Stimulus Condition	Kind of Stimulus	Rating Condition				All Groups	Control (EO)
		R4	R3	R2	R1		
SS	Extra	0.08	-1.33	-0.42	-0.42	-0.52	0.42
	Standard	0.00	0.00	-0.08	-0.25	-0.08	0.50
DS	Extra	0.33	-0.33	-0.08	-0.50	-0.15	
	Standard	-0.50	-0.50	-1.08	0.00	-0.52	
All Groups	Extra	0.21	-0.83	-0.25	-0.46	-0.34	
	Standard	-0.25	-0.25	-0.58	-0.13	-0.30	

other stimulus; (b) a score of 1 was assigned if S reinforced several stimuli, among which were both instances of the CS; (c) a score of 0 was assigned if S reinforced several stimuli including an instance of the CS, failed to reinforce either occurrence of the CS, reinforced every spin, or failed to reinforce any spin. These scores reflect Ss knowledge of the correct behavioral hypotheses. The weighted means for all groups are presented in Table XI.

Responses to the modified Dulany interview were also made into composite awareness scores to reflect behavioral hypotheses. The scores (summarized in Table XI) were determined as follows: (a) A score of 2 was registered if, according to his answers on questions 1, 2, and/or 3 of the interview (see p. 70), S reported the correct response-reinforcement contingency. (b) If S reported a correlated, but incorrect, hypothesis he received a score of 1. An example of such a hypothesis was reported by one child, for whom the Greek letter Σ was the critical stimulus. He reported that a marble was delivered each time δ or Σ or Δ occurred. (c) If S reported an incorrect noncorrelated hypothesis or no hypothesis concerning the response contingency, a score of 0 was assigned.

Responses to questions 1 and 2 indicated that all Ss (for whom a CS was reinforced) had the correct reinforcement hypothesis; i.e., they all reported that a marble sometimes emerged when the arrow stopped. However, two children had incorrect reinforcement hypotheses when they played the role of E.[5] All data sheets of

---

[5] The first child dispensed reinforcements while the arrow was still spinning while the second child delivered no reinforcements. However, the subsequent interview revealed that this second S had

TABLE XI

WEIGHTED MEANS FOR POSTEXPERIMENTAL AWARENESS MEASURES (SCORES OF 0, 1, 2) COMPARED WITH MEANS OF FINAL PREFERENCE SCORE

Stimulus Condition	Kind of Score	Rating Condition				All Groups
		R4	R3	R2	R1	
SS	Role	1.42	1.42	1.42	1.50	1.44
	Interview	1.50	1.50	1.50	1.67	1.54
	Preference	2.92	2.18	1.25	1.08	1.86
DS	Role	0.50	0.75	0.58	0.50	0.58
	Interview	1.00	0.83	0.83	0.75	0.85
	Preference	0.00	0.92	0.50	0.50	0.48
All Groups	Role	0.96	1.09	1.00	1.00	1.01
	Interview	1.25	1.17	1.17	1.21	1.20
	Preference	1.46	1.55	0.87	0.79	1.17

individual Ss whether from the role-playing procedure or from the interview procedure were scored "blind."

There was close agreement between the scores based on the two awareness measures. A product moment correlation of .90 was obtained for these scores, supporting the assumption that the behavioral data obtained through reversed reinforcement roles are complementary to those obtained by the Dulany interview procedure. Weighted means (Table XI) for the interview were slightly, but consistently higher than role means. Since the role reversal preceded the interview it may have had a carryover effect. Nevertheless, the results from the two procedures appear to be sufficiently promising to be worthy of further refinement for future studies with young children where measures of awareness are required.

Question 4 and 5 of the interview were designed to investigate any verbalizations that these Ss were able to make concerning demand characteristics. Many Ss answered "yes" to question 5c (see p. 70), but the validity of such reports is suspect in light of the sparse number of replies obtained in probes for items 4 and 5. In only one case did an S report of his own volition (question 4) that he thought that E wanted him to assign "good words to the picture that got marbles."

---

an incorrect behavioral hypothesis rather than an incorrect reinforcement hypothesis. Thus, on being questioned the S indicated that she received a marble each time the arrow came to rest on the Ω symbol, which was correct. When asked why she did not push the button when E's spin stopped on Ω, she said, "It is long hair; you have short hair, and there aren't any short hair pictures." Her behavioral hypothesis was that only people with long hair should receive a reinforcement when the arrow pointed to omega since its shape resembles the outline of a girl's hair.

An examination of the data showing the relationship between "awareness" and preference (summarized in Table XI) indicates gross support for the hypothesis that Ss who had correct reinforcement and behavioral hypotheses were "conditioned." Thus, the reader will note in that table the correspondence between the mean role taking score, the mean interview score, and the mean adjective rating for the final experimental day. Role, interview, and preference means for the SS condition were respectively higher than the corresponding scores for the DS treatments, but differences between SS groups cannot be explained on this basis since all SS groups evidenced the same level of awareness. These may also be compared with the analysis summarized in Table VII in which the SS and DS ratings were shown to be different. In the present analysis, SS role and interview means were also found to differ significantly from DS role and interview means ($t(94) = 4.89$, $p < .01$ and $t(94) = 3.92$, $p < .01$) respectively. Thus, awareness, as revealed by postexperimental measures, was correlated with "conditioning" of preference; however, it was not correlated with differences among SS groups.

CHAPTER VI

DISCUSSION

The results of the experiment are in essential agreement with the hypotheses. In general, these results are discussed with respect to the roles of reinforcement, repeated evaluations, and awareness in semantic conditioning.

Reinforcement Procedures in the Development of Preferences

The evidence clearly supports the preliminary assumption (Hypothesis I) that preferences can be "conditioned" by associating stimuli with secondary reinforcers. This is a principle which has recently been established but was necessary to replicate in the present study to assure that the effect could be obtained before the subsequent hypotheses were examined. The slope and trend of the performance curve for Group SS-R4 produced an identical replication of the Nunnally, Stevens, and Hall (1965) experiment in which a comparable procedure was employed, attesting to the reliability of the effect. Accordingly, these data warrant the conclusion that when a neutral stimulus is presented contiguously with secondary reinforcers, for example marbles or trinkets, first grade children's evaluations of this stimulus increase across rating days which are alternated with training or conditioning days. This conclusion is consistent with those from previous studies (Schroder, 1956;

Nunnally, Stevens, & Hall, 1965; Nunnally, Duchnowski, & Parker, 1965; Parker & Nunnally, 1966; Nunnally, Knott, & Duchnowski, 1967; Knott, Nunnally, & Duchnowski, 1967; Faw & Nunnally, 1967; Hall, 1967; Nunnally, Duchnowski, & Knott, 1967; Nunnally & Faw, 1968).

Simple *contiguity* between a reinforcing stimulus and a neutral stimulus has been demonstrated (e.g., by Schroder, 1956) to lead, over trials, to preference for the formerly neutral stimulus. Since children tend to respond with consistency even in the absence of reinforcement, as the Coefficient of Concordance analysis indicated for the SS-EO group, the contiguity between CS and reinforcement may be thought of as directing *S*s' tendencies for consistent responding toward the critical stimulus. With regard to an experiment in which *S* was reinforced by *E* saying "right" for the response he wanted to condition and "wrong" for all others, Thorndike (1935) found an additional cumulative effect of reinforcement. He states that even when the subject has just been informed that the last preceding connection is wrong, the satisfaction of hearing "four right of the last eight . . . you're doing fine" or the likes does demonstrably strengthen the correct connection. Similarly, in addition to contiguous association of critical and reinforcing stimulus in the present investigation, the summary or cumulative reinforcing effect of pasting stars upon his paper and knowing that he is making progress toward earning a toy may positively affect *S*'s preference for the CS.

Studies on *affective contrast* (Helson, 1964) tend to indicate that a given stimulus may be judged as pleasant or unpleasant depending on whether an S has just been seeing other stimuli which are relatively unpleasant or pleasant, respectively, in relation to the one being judged at the moment. Thus, if S learns that most of the stimuli never pay-off, his preference for the one that does pay-off may be enhanced relative to the ratio of pay-off to non-pay-off stimuli (trials).

The cumulative effects of contingent presentation and affective contrast may be posited to give emphasis to the critical stimulus. If, over the course of the experiment, S comes to contemplate and look at the critical stimulus more than nonrewarded stimuli as selective attention studies by Nunnally and associates indicate, then his effective exposure to that stimulus is greater. Accordingly, one may argue, as Zajonc's (1968) experiments do, that "mere repeated exposure of the individual to a stimulus object enhances his attitude toward it" (Zajonc, 1968, p. 1). This simply means that, although absolute exposure to all standard stimuli is precisely the same, in effect, the emphasis given to the critical stimulus causes S to attend to that stimulus more than to the others. We can think of reinforcement, then, as increasing selective or effective exposure and thus increasing preference through the additional mechanism of greater exposure to the critical stimulus.

Thus, we see that the well established role of reinforcement is strongly supported, as expected, by the present experiment. The reasons for the strengthening effect of reinforcements have been variously interpreted to conform with the theoretical predilection of individual researchers. However, it should be recalled that one of the primary purposes of the present study was to determine what, if any, effect repeated ratings have on preference.

Ratings and Preferences

The results of the Jonckheere (1964) test clearly upheld the hypothesis that preference for the critical stimulus in the SS groups is also (in addition to reinforcement) a direct function of the number of ratings. Within the framework of Orne's demand characteristics hypothesis or of Dulany's propositional verbal control, it can be hypothesized that additional ratings function to increase Ss' knowledge (whether with awareness or not) of what the E requires. This knowledge is obviously more readily gained under the SS conditions than under the DS conditions. Nevertheless, all of the 14 points in Figure 5a and all but one of the 14 points in Figure 5b were in the ordinal position predicted. In keeping with general principles of acquisition, the learning situation which provides for the consistency of experience, existent in the SS condition, has a decided advantage in its potentiality for Ss' acquisition of the principle (whatever the relationship) to be transferred. As a consequence, the effects of learning this principle on preference is pronounced and statistically significant. On the other hand, the less consistent experience provided by the DS conditions permits correspondingly less opportunity for acquisition of the principle due to such factors as the new or substitute contingencies that must be learned. The potentiality for learning and transfer is also correspondingly lessened. Nevertheless, though the effect of ratings on preference remains clearly apparent, its significance is not statistically demonstrable in the DS conditions.

As one might suspect there may be a number of principles that can be learned and transferred to affect preferences. Among these are the actual or implied intent of the E, some consistent classification based on an attribute that occurs to S, a clarification of the rating procedure, and the stimulus that pays-off. Thus, the number of evaluations of the stimulus influences the degree of polarity of such evaluations. Conversely, if repeated ratings are experimentally partialed out from the factors that contribute to the conditioning of preferences then a slight reduction in polarity occurs.

Carrying Zajonc's premise, that mere frequency of exposure enhances preference, a step further, a frequency mechanism similar to Ekstrand, Wallace, and Underwood's (1966) theory of verbal discrimination learning may be hypothesized. Accordingly, each time S effectively attends to the reinforced stimulus, one unit of habit strength is accumulated. Attending during the rating sessions as well as assigning positive adjectives to the stimulus can be thought of as adding to habit strength, while assignment of a negative adjective subtracts a unit of habit strength. Whether or not such units are of identical size is not of practical importance to the present discussion. But the important consideration is that a greater number of habit strength units are built up for groups who rate the stimuli more often.

It can also be argued that conditioning of preference value through secondary reinforcement with humans, to the extent that Ss are sensitive to experimental demands, is a special case of transfer. Take for example the case of a child who is confronted with the stimuli on the final day in the SS conditions. He has participated in six previous experimental sessions, and, with the exception of the extra stimulus, now views stimuli seen on three previous conditioning days. In the conditioning sessions he was reinforced each time his spin stopped at the critical stimulus. During the alternate three sessions he has a history of rating activity, filler activity, or some combination of the two activities, depending on the experimental condition in which he participated. Then, in rating the stimuli on the seventh day, Ss who have had rating experiences previously tend to evaluate the stimulus more positively than Ss who have never participated in the procedure. This superiority may be the product of nonspecific transfer factors and can be interpreted according to learning set theory (Harlow, 1949; Di Vesta & Walls, 1966b). In this case an explanation would necessarily presuppose that nonassociative factors in transfer were influencing ratings with increasing influence across problems. To the extent that S learns "how" he should evaluate the stimuli he has seen before, his learning set will correspond closely to demands implicit in the task of rating the reinforced stimulus in a favorable manner.

It is probable that most Ss do form hypotheses about the nature of the task requirements. As Farber states, "Subjects may not know exactly what is going on in an experiment or, for that matter, in a therapeutic session, but very few have no ideas at all. They may be mistaken, or they may be concerned with irrelevant . . . ." (Farber, 1963, p. 196) matters, but perception of experimental demand characteristics is a very real phenomenon (Orne, 1962). Mediational association (conscious or unconscious) of the verbal rating procedure employed previously, with his acquired learning set (whether correct or incorrect) may produce subsequent responding consonant with his perception of any given phase of the experimental situation.

The situation may even be somewhat analogous to the reception paradigm in concept attainment (Bruner, Goodnow, & Austin, 1956) in which S receives successive presentations of neutral stimuli and tries to identify the relevant rules and attributes of the concept that E has in mind. While this appears a rather extreme cognitive interpretation, it represents the process that a sophisticated S may employ. For example, where semantic conditioning is readily achieved with as few as five conditioning trials as in the study cited earlier (Nunnally, Knott, & Duchnowski, 1967), it appears likely that the discrimination of the critical stimulus is an easily mastered task. That study differed in a number of ways (e.g., only three stimuli, one punished, one rewarded, and one neither punished nor rewarded) from the present investigation.

Changes in preferences are undoubtedly due primarily to the influence of reinforcements in the present experiment. Nevertheless, ratings and experience unquestionably also have their effect. The present investigator would even entertain the hypothesis that stronger effects due to "demand characteristics" learned as a learning set through repeated ratings in the DS conditions might have been obtained had greater numbers of "problems" (rating days) been used. This is a consideration of theoretical importance that may be worthy of further investigation in view of the suggestive results obtained in the present investigation.

Awareness and Preferences

An additional purpose of the experiment was to determine the functional relationship between the "conditioning" of preferences and awareness. Granted that a relation between awareness and a S's rating behavior exists, there remains the perennial question of whether awareness is the product of modified behavior or whether the behavior is a consequence of the awareness, and further, whether the question itself is even a meaningful one. As Lyons (1965) comments, "To the nonscientist, the answer is obvious. He would insist that in the ordinary case a person first becomes able to do something, then decides to do it, then does it, and finally says that he does it. But what is obvious to the nonscientist is neither obvious nor necessary for experimental psychologists, and so on this issue one often finds them divided" (Lyons, 1965, p. 260).

While the specification of a cuase-and-effect sequence is beyond the scope of the present investigation, the results indicate that Ss who properly reinforce the rewarded stimulus in the role taking session and report the response-reinforcement contingency in the interview, also evaluate the CS more positively. Generally, "conditioning" is not obtained unless Ss are aware and unless they have acquired correct reinforcement and behavioral hypotheses. It should be recalled here that approximately three-fourths of the Ss in each SS group, but only one-fourth in DS groups were cognisant of this contingency. This is not a surprising result in view of the widely different number of reinforcements administered for a given critical stimulus in the two sets of groups. However, since the level of awareness in each SS group was the same, differences between those groups may not be explained in terms of awareness. There is the possibility, in this regard, that the present awareness measures may not have been sensitive enough to measure knowledge of the task beyond the response-reinforcement contingency. Additionally, any statement regarding the relationship between awareness and learning can be no more definitive on the basis of present results than from previous studies because the number of reinforcements is confounded with level of awareness. At best only an R-R relationship between level of awareness and level of preference can be posited.

With respect to the implicit hypothesis concerning the complementary relationship between reversal of reinforcement roles and interview procedures, the findings are clear. The reversal situation in which S manipulates the reinforcement in an attempt to duplicate earlier proceedings is a behavioral manifestation of constructs which the Dulany interview is designed to measure. This technique would be of particular value for experiments in which human learning is measured through operant responding or in a classical conditioning situation which allows reinforcement to be delivered manually or orally. The refinement of this technique might contribute to the reduction of the awareness controversy by providing an empirical base for inference rather than to rely on the person's ability to verbalize his experience.

In conclusion, the results of this experiment attest to the pervasive influence of reinforcement. Reinforcement was found to be the dominant influence in the development of preferences, and thereby, the findings of Nunnally, Stevens, and Hall (1965) have been supported. However, the present results also show that repeated ratings of stimuli is also a contributing factor and, in extended experience as might occur in everyday life, might even be a substantial one, though never to the extent that the influence of reinforcement would be overshadowed.

CHAPTER VII

SUMMARY

The present experiment was designed to determine the effects of frequency of reinforcement, and repeated evaluation of stimuli on inferred cognitive correlates in semantic conditioning, with particular reference to the conditioning of preferences. Previous experimental results have dealt exclusively with the effect of contiguous relationships between stimuli and reinforcement in this process.

The area of investigation treated in this dissertation is in the domain of studies on the law of effect. Experiments which have followed and built upon Thorndike's early work on the law of effect led to differences of opinion regarding the role of awareness and cognitive factors on learning. Verbal operant conditioning is one area in which the question of whether Ss' verbal behavior can be modified without their knowledge of that change has been challenged. Earlier studies of verbal conditioning may be divided into two categories in this respect. Some studies failed to report an assessment of awareness of the response-reinforcement contingency, although the majority of investigations reported their Ss to be unaware of this contingency. Recent works, in which a much greater effort has been devoted to precise measurement of awareness, report almost universal awareness of the response-reinforcement contingency by Ss who were conditioned. This striking revision in outlook has been a product of refinement and partial resolution of the criterion problem. Nevertheless, even today the Dulany-Locke technique which is among the most widely employed for measuring awareness, appeals to _S_'s capacity for introspection, a procedure which is viewed with disfavor by many behaviorists. Similar difficulties have been encountered in classical conditioning of semantic meaning. While some investigators (e.g., Staats & Staats, 1957) report no evidence that _S_s become aware of the CS-UCS contingency, others (e.g., Maltzman, 1966) indicate that _S_s often are aware of the CS-UCS relationship and, additionally, that such _S_s "condition" with reliable superiority.

It is feasible that "conditioning" in verbal and/or semantic conditioning may be related to _S_'s perception of experimental demand characteristics (Orne, 1962). Further, the possibility exists that the constructs implicit in the conditioning operation and perceived demands are related through unconscious or conscious mediation.

With respect to conditioning of meaning in children's learning, the general conclusion is that rewarded stimuli are preferred by children more than punished or neutral stimuli. Preferential verbal evaluation and selective attention have been viewed as being more adequately accounted for by the automatic action of reinforcement than by cognitive factors. Within these series of investigations, Nunnally, Stevens, and Hall (1965) found a linear increase in preference for a critical stimulus over four rating days. Examination of the results of that experiment revealed that the possible effects of conditioning (reinforcements) and cognitive change, or awareness of demand characteristics due to number of ratings, on the development of preferences if, indeed, cognitive factors did have an effect, could not be clearly separated.

Accordingly, an experiment was designed to test the hypotheses that: (a) The development of preferences can be facilitated by associating stimuli with secondary reinforcers. (b) The magnitude of the preference for a given conditioned stimulus is a function of the number of reinforcements associated with the stimulus. (c) The magnitude of the preference is directly related to the number of ratings of a given stimulus made by S. (d) Ratings of conditioned stimuli in which all stimuli are varied from one conditioning session to the next will increase in polarization over the course of the experiment due, primarily, to the effects of rating. It was predicted that this effect should be weaker than that in (c), above, to the extent that reinforcement had an effect. However, the alternative possibility that preferences were developed exclusively as a function of reinforcements was not ruled out.

The experimental design used to investigate these assumptions consisted of four levels of rating schedules crossed with two levels of stimulus conditions. In addition, a group of Ss (Condition SS-NO) which had experience with the stimuli but no reinforcements was separate from the 2 x 4 design. There were 12 first grade Ss in each of these nine conditions.

In one variation of the stimulus conditions the stimuli were always the same from one experimental session to the next (Condition

SS); in the other the stimuli differed from one conditioning session to the next (Condition DS). Of the seven consecutive school days in which the experiment was conducted, Ss in one group (Group R4) rated stimuli on four days (days 1, 3, 5, 7); another group (Group R3) rated stimuli on three days (days 3, 5, 7); a third group (Group R2) rated stimuli on two days (days 5, 7); and a fourth group (Group R1) rated stimuli on one day (day 7). Ss in Groups R3, R2, and R1 were administered filler tasks on days in which ratings were not obtained. The control group (SS-R0) rated stimuli on all four rating days but received no reinforcements on conditioning days. Thus, the eight experimental groups and one control may be represented as SS-R4, SS-R3, SS-R2, SS-R1, DS-R4, DS-R3, DS-R2, DS-R1, and SS-R0.

Conditioning trials were administered to all experimental Ss on days 2, 4, and 6. A conditioning session consisted of S spinning the arrow of the wheel-spin apparatus until 20 reinforcements had been received. E reinforced S with marbles by an automatic dispenser approximately 0.5 second after the pointer came to rest on the preselected critical stimulus. The marbles were later exchanged for a toy.

The stimuli for the conditioning and rating tasks were four sets of upper and lower case Greek letters with seven letters in a set. Each set consisted of six standard stimuli and an extra stimulus which did not appear on the conditioning apparatus but was used in rating trials.

The measure of preference (rating) was based on the procedure of assigning 14 adjectives (seven positive evaluative and seven negative evaluative) to the rating stimuli. Within a rating session, the six appropriate wooden stimulus blocks were set before S. He was instructed to point to the stimulus which he felt was best described by a given adjective. A "preference score" was obtained by subtracting the number of negative adjectives assigned to the reinforced (critical) stimulus from the number of positive adjectives assigned to that stimulus.

Postexperimental measures of awareness were obtained by reversal of reinforcement roles and by interview data to provide information about S's reinforcement and behavioral hypotheses. The former procedure involved a reversal of roles in which S's task was to correctly reinforce E as E spun the arrow of the conditioning apparatus. In the latter procedure a series of questions, adapted from Dulany's (1962) theoretical framework, was employed as a second basis for inference concerning S's reinforcement and behavioral hypotheses as they relate to propositional verbal control.

A replication of the Nunnally et al. (1965) experiment was produced by the comparable condition (SS-R4) in the present study. Analysis of variance results clearly showed that the combination of conditions and variables in the Nunnally studies provided reliably demonstrable increases in preferences for stimuli associated with reinforcements. Though the demonstration of this effect is not a unique finding of the present study, it is basic to the test of subsequent hypotheses. The interpretation proposed to account for the effects of reinforcement was as follows: Simple contiguity (Thorndike, 1932) between a reinforcing and a neutral stimulus (which directs consistent responding to the CS) and affective contrast (Helson, 1964) between reinforced and nonreinforced stimuli combine to give emphasis to the critical stimulus. Such emphasis increases the frequency of effective exposure (Zajonc, 1968) of that stimulus as compared to nonrewarded stimuli and concomitantly increases _S_'s preference for it. Thus, the "satisfying" effect of the contiguity between stimulus and reward is supplemented by the mechanism of increased effective exposure.

The results of the Jonckheere (1954) test upheld the hypothesis (Hypothesis II) that preference for the critical stimulus in the SS set of groups is, in addition to reinforcement, a direct function of the number of ratings. An increase of awareness of the purposes and procedures of the experiment along with the acquisition of a learning set were suggested as possible consequences of the rating procedure, thereby leading to an increase in _Ss_' knowledge (whether with awareness or not) of experimental demands.

This "rating" effect was stronger for SS than DS groups, as can be expected on the basis of such factors as reinforcement, mere frequency of effective exposure, and habit strength. A frequency theory was proposed which would predict that the more often a given stimulus is effectively exposed or rated positively, the more habit strength units are acquired. Thus, Hypothesis III, which stated that ratings of conditioned stimuli in which all stimuli are varied from one conditioning session to the next will increase in polarization over the course of the experiment due, primarily, to the effects of rating, was not strongly supported by the results. However, the general configuration of the DS curves (13 of 14 points in the predicted order) suggested that in real life situations or in experiments of longer duration, a more clear-cut "rating" effect, even with different stimuli, may be demonstrable.

The results also suggest that the interview complements the role reversal procedure in obtaining awareness measures. The effect of number of reinforcements on awareness as measured by both procedures was apparent. Thus, approximately three fourths of the Ss in each SS group but only one fourth in each DS group were aware of the response-reinforcement contingency. Both measures clearly differentiated between SS and DS groups which are differentially delimited according to degree of reinforcement for a given critical stimulus. Awareness appears to be a function of reinforcement, and further, behavioral hypotheses and cognizance of demand characteristics are related to awareness. However, no really definitive causal relationship between awareness and performance can be made, and awareness, as measured in the present study, cannot account for SS rating differences. The role exchange procedure, through provision of a directly observable behavior base for inference and with improved techniques, may permit further refinement in the measurement of degree of awareness.

In general, the findings in the present study tend to argue for (a) an awareness of the response-reinforcement contingency which is concomitant with correct reinforcement and behavioral hypotheses and "conditioning"; (b) conscious or unconscious perception of experimental demands mediated by an acquired learning set; (c) the strength of the dominant effects of reinforcement on preferences. Experimenters should be cognisant of the possible effects of situations which call for repeated evaluations and provide controls for such techniques. The finding that the number of evaluations of an object that a person has made influences his subsequent evaluation is an interesting phenomenon and a psychological problem in its own right. It has implications in the broader context of human experience as well as in the laboratory.

REFERENCES

Adams, J. K. Laboratory studies of behavior without awareness. *Psychological Bulletin*, 1957, *54*, 383-405.

Atkinson, R. C., Bower, G. H., & Crothers, E. J. *An introduction to mathematical learning theory*. New York: Wiley, 1965.

Box, G. E. P. Some theorems on quadratic forms applied in the study of analysis of variance problems. *Annals of Mathematical Statistics*, 1954, *25*, 290-302, 484-498.

Bruner, J. S., Goodnow, Jacqueline J., & Austin, G. A. *A study of thinking*. New York: Wiley, 1956.

Bulgarella, R. A. The role of awareness in verbal learning. *Dissertation Abstracts*, 1966, *26*, 4826.

Bugelski, B. R., & Scharlock, D. P. An experimental demonstration of unconscious mediated association. *Journal of Experimental Psychology*, 1952, *44*, 334-338.

Buss, A. H., & Gerjuoy, Irma R. Verbal conditioning and anxiety. *Journal of Abnormal and Social Psychology*, 1958, *57*, 249-250.

Campbell, D. T., & Stanley, J. C. Experimental and quasi-experimental designs for research on teaching. In N. L. Gage, (Ed.), *Handbook of research on teaching*. Chicago: Rand-McNally, 1963. Pp. 171-246.

Carroll, J. *Language and thought*. Englewood Cliffs, N. J.: Prentice-Hall, 1964.

Chatterjee, B. B., & Eriksen, C. W. Conditioning and generalization of GSR as a function of awareness. *Journal of Abnormal and Social Psychology*, 1960, *60*, 396-403.

Cofer, C. N., & Foley, J. P. Mediated generalization and the interpretation of verbal behavior: I. Prologemena. *Psychological Review*, 1942, *49*, 513-540.

Cohen, B. D., Kalish, H. I., Thurston, J. R., & Cohen E. Experimental manipulation of verbal behavior. *Journal of Experimental Psychology*, 1954, *47*, 106-110.

Cohen, B. H. Role of awareness in meaning established by classical conditioning. *Journal of Experimental Psychology*, 1964, *67*, 373-378.

Coover, J. E. Experiments in psychical research. *Psychical Research*, *Monograph No. 1*. Stanford U. Press, 1917. Cited by Adams, 1957.

Cowen, E. L., & Beier, E. G. The influence of threat expectancy on perception. *Journal of Personality*, 1950, *19*, 85-94.

Davis, K. G., & Hess, H. F. The effectiveness of concepts at various levels of awareness. *Journal of Experimental Psychology*, 1962, *63*, 62-67.

Dean, S. J., & Hiesinger, L. Operant level, awareness, and the Greenspoon effect. *Psychological Reports*, 1964, *15*, 931-938.

De Nike, L. D. The temporal relationship between awareness and performance in verbal conditioning. *Journal of Experimental Psychology*, 1964, *68*, 521-529.

Diven, K. Certain determinants in the conditioning of anxiety reactions. *Journal of Psychology*, 1937, *3*, 291-308.

Di Vesta, F. J. Effects of mediated generalization on the development of children's preferences for figures. *Child Development*, 1962, *33*, 209-220.

Di Vesta, F. J. The measurement of children's meaning. *Pedagogisk Forskining*, 1966, 150-168.

Di Vesta, F. J., & Blake, Kathryn. The effects of instructional 'sets' on learning and transfer. *American Journal of Psychology*, 1959, *72*, 57-67.

Di Vesta, F. J., & Stover, D. O. The semantic mediation of evaluative meaning. *Journal of Experimental Psychology*, 1962, *64*, 467-475.

Di Vesta, F. J., & Walls, R. T. Response selection as a function of instructions and motivation under nonreinforcement conditions. *Journal of Experimental Psychology*, 1967, *73*, 365-373. (a)

Di Vesta, F. J., & Walls, R. T. Transfer of object-function in problem solving. *American Educational Research Journal*, 1967, *4*, 207-215. (b)

Di Vesta, F. J., & Walls, R. T. Transfer of solution rules in problem solving. *Journal of Educational Psychology*, 1967, *58*, 319-326. (c)

Dixon, P. W., & Oakes, W. F. Effect of intertrial activity on the relationship between awareness and verbal operant conditioning. Journal of Experimental Psychology, 1965, 69, 152-157.

Dulany, D. E. Hypotheses and habits in verbal "operant conditioning." Journal of Abnormal and Social Psychology, 1961, 63, 251-263.

Dulany, D. E. The place of hypotheses and intentions: an analysis of verbal control in verbal conditioning. Journal of Personality, 1962, 30, (2, Suppl. No. 6), 102-129.

Eisman, B. S. Attitude formation: The development of a color-preference response through mediated generalization. Journal of Abnormal and Social Psychology, 1955, 50, 321-326.

Ekstrand, B. R., Wallace, W. P., & Underwood, B. J. A frequency theory of verbal discrimination learning. Psychological Review, 1966, 73, 566-578.

Eriksen, C. W. Discrimination and learning without awareness: a methodological survey and evaluation. Psychological Review, 1960, 67, 279-300.

Eriksen, C. W., & Doroz, L. Role of awareness in learning and use of correlated extraenous cues on perceptual tasks. Journal of Experimental Psychology, 1963, 66, 601-608.

Eriksen, C. W., & Kuethe, J. L. Avoidance conditioning of verbal behavior without awareness: a paradigm of repression. Journal of Abnormal and Social Psychology, 1956, 53, 203-209.

Essman, W. B. Awareness of reinforcement in learning without awareness. Psychological Reports, 1957, 3, 399-400.

Farber, I. E. The things people say to themselves. American Psychologist, 1963, 18, 185-197.

Finley, J. R., & Staats, A. W. Evaluative meaning words as reinforcing stimuli. Journal of Verbal Learning and Verbal Behavior, 1967, 6, 193-197.

Gerstein, A. I. The effect of reinforcement schedules on meaning generalization and on awareness of the purpose of the experiment. Journal of Personality, 1961, 29, 350-362.

Golin, S. Recall as a function of awareness of CS-UCS contingency. Psychological Reports, 1964, 15, 919-923.

Greenspoon, J. The effect of a verbal stimulus as a reinforcement. Proceedings of the Indiana Academy of Science, 1950, 59, 287.

Greenspoon, J. The reinforcing effect of two spoken sounds on the frequency of two responses. *American Journal of Psychology*, 1955, *68*, 409-416.

Grings, W. W. Verbal-perceptual factors in conditioning of autonomic responses. In W. F. Prokasy (Ed.), *Classical conditioning: a symposium*. New York: Appleton-Century-Crofts, 1965, 71-89.

Hall, G. F. Association of neutral objects with rewards: Persistence of effect upon verbal evaluation. *Journal of Verbal Learning and Verbal Behavior*, 1967, *6*, 528.

Harlow, H. F. The formation of learning sets. *Psychological Review*, 1949, *56*, 51-65.

Hefferline, R. F., Keenan, B., & Harford, R. A. Escape and avoidance conditioning in human *Ss* without their observation of the response. *Science*, 1959, *130*, 1338-1339.

Helper, M. M. A study of pictorial scales with retarded persons. Progress Report. USPHS Grant No. MH-10793, Unpublished Manuscript, Nebraska Psychiatric Institute, 1965.

Helson, H. *Adaptation-level theory*. New York: Harper & Row, 1964.

Hirsch, J. Learning without awareness and extinction following awareness as a function of reinforcement. *Journal of Experimental Psychology*, 1957, *54*, 218-224.

Horton, D. L. The effects of meaningfulness, awareness, and type of design in verbal mediation. *Journal of Verbal Learning and Verbal Behavior*, 1964, *3*, 187-194.

Horton, D. L., & Kjeldergaard, P. K. An experimental analysis of associative factors in mediated generalization. *Psychological Monographs*, 1961, *75*, (11, Whole No. 515).

Hull, C. L. The goal gradient hypothesis applied to some 'field-force' problems in the behavior of young children. *Psychological Review*, 1938, *45*, 271-299.

Irwin, F. W., Kaufman, K., Prior, G., & Weaver, H. B. On "Learning without awareness of what is being learned." *Journal of Experimental Psychology*, 1934, *17*, 823-827.

Janke, P. K. The effects of a 'compound set' on learning efficiency. *American Journal of Psychology*, 1962, *75*, 445-451.

Jonckheere, A. R. A distribution-free k-sample test against ordered alteratives. *Biometrika*, 1954, *41*, 133-145.

Kanfer, F. H., & McBrearty, J. F. Verbal conditioning: Discrimination and awareness. *Journal of Psychology*, 1961, *52*, 115-124.

Katkin, E. S., Risk, R. T., & Spielberger, C. D. The effects of experimenter status and subject awareness on verbal conditioning. *Journal of Experimental Research in Personality*, 1966, *1*, 153-160.

Kendler, Tracy S., Kendler, H. H., & Wells, Doris. Reversal and nonreversal shifts in nursery school children. *Journal of Comparative and Physiological Psychology*, 1960, *53*, 83-88.

Kimbel, G. A. Classical conditioning and the problem of awareness. *Journal of Personality*, 1962, *30*, (2, Suppl. No. 6), 27-45.

Klein, B., & Wiener, M. Awareness in the "learning without awareness" paradigm. *Journal of Experimental Research in Personality*, 1966, *1*, 145-152.

Knott, P. D., Nunnally, J. C., & Duchnowski, A. J. Effects of frustration on primary and conditioned incentive value. *Journal of Experimental Research in Personality*, 1967, *2*, 140-149.

Krasner, L. Studies in the conditioning of verbal behavior. *Psychological Bulletin*, 1958, *55*, 148-171.

Krieckhaus, E. E., & Eriksen, C. W. A study of awareness and its effects on learning and generalization. *Journal of Personality*, 1960, *28*, 503-517.

Lacey, J. I. & Smith, R. L. Conditioning and generalization of unconscious anxiety. *Science*, 1954, *120*, 1045-1052.

Lanyon, R. I. Verbal conditioning and awareness in a sentence construction task. *American Journal of Psychology*, 1964, *77*, 472-475.

Lashley, K. S. *Brain mechanisms and intelligence*. Chicago: University Press, 1929.

Leventhal, A. M. The effects of diagnostic category and reinforcer on learning without awareness. *Journal of Abnormal and Social Psychology*, 1959, *59*, 162-166.

Levin, S. M. The effects of awareness on verbal conditioning. *Journal of Experimental Psychology*, 1961, *61*, 67-75.

Locke, E. A. Logical relationships of questions to issues in verbal conditioning studies. *Psychological Reports*, 1966, *19*, 291-298.

Lyons, J. *A primer of experimental psychology*. New York: Harper & Row, 1965.

Maltzman, I. Awareness: Cognitive psychology vs. behaviorism. *Journal of Experimental Research in Personality*, 1966, *1*, 161-165.

Mandler, G., & Kaplan, W. K. Subjective evaluation and reinforcing effect of a verbal stimulus. *Science*, 1956, *124*, 582-583.

Mednick, S., & Freedman, J. L. Stimulus generalization. *Psychological Bulletin*, 1960, *57*, 169-200.

Miller, J. G. *Unconsciousness*. New York: Wiley, 1942. Cited by Adams, 1957.

Miller, N. E., & Dollard, J. *Social learning and imitation*. New Haven, Connecticut: Yale University Press, 1941.

Miron, M. S., & Osgood, C. E. Language behavior: the multivariate structure of qualification. In R. B. Cattell, (Ed.) *Handbook of multivariate experimental psychology*. Chicago: Rand McNally, 1966, pp. 790-819.

Mowrer, O. H. The psychologist looks at language. *American Psychologist*, 1954, *9*, 660-694.

Newhall, S. M., & Sears, R. R. Conditioning finger reaction to visual stimuli near the absolute threshold. *Comparative Psychological Monograph*, 1933, *9*, No. 43.

Nunnally, J. C., Duchnowski, A. J., & Knott, P. W. Association of neutral objects with rewards: Effects of massed versus distributed practice, delay of testing, age, and sex. *Journal of Experimental Child Psychology*, 1967, *5*, 152-163.

Nunnally, J. C., Duchnowski, A. J., & Parker, R. K. Association of neutral objects with rewards: Effect on verbal evaluation, reward expectancy, and selective attention. *Journal of Personality & Social Psychology*, 1965, *1*, 270-274.

Nunnally, J. C., & Faw, T. T. The acquisition of conditioning reward value in discrimination learning. *Child Development*, 1968, *39*, 159-166.

Nunnally, J. C., Knott, P. D., & Duchnowski, A. J. Association of neutral objects with rewards: Effects of different numbers of conditioning trials and of anticipated reward versus actual reward. *Journal of Experimental Child Psychology*, 1967, *5*, 249-262.

Nunnally, J. C., Stevens, D. A., & Hall, G. F. Association of neutral objects with rewards: Effects on verbal evaluation and eye movements. *Journal of Experimental Child Psychology*, 1965, *2*, 44-57.

Orne, M. T. The nature of hypnosis: Artifact and essence. *Journal of Abnormal and Social Psychology*, 1959, *58*, 277-299.

Orne, M. T. On the social psychology of the psychological experiment: With particular reference to demand characteristics and their implications. *American Psychologist*, 1962, *17*, 776-783.

Orne, M. T., & Evans, F. J. Inadvertent termination of hypnosis with hypnotized and simulating subjects. *International Journal of Clinical and Experimental Hypnosis*, 1966, *14*, 61-78.

Osgood, C. E. The nature and measurement of meaning. *Psychological* Bulletin, 1952, *49*, 197-237.

Osgood, C. E., Suci, G. J., & Tannenbaum, P. H. *The measurement of meaning*. Urbana: University of Illinois Press, 1957.

Osipow, S. H. The effects of verbal mediation on the modification of children's attitudes. *Journal of Educational Psychology*, 1960, *51*, 199-207.

Parker, R. K., & Nunnally, J. C. Association of neutral objects with rewards: Effects of reward schedules on reward expectancy, verbal evaluation, and selective attention. *Journal of Experimental Child Psychology*, 1966, *3*, 324-332.

Philbrick, E. B., & Postman, L. A further analysis of 'learning without awarenes.' *American Journal of Psychology*, 1955, *68*, 417-424.

Pollio, H. R. Word association as a function of conditioned meaning. *Journal of Experimental Psychology*, 1963, *66*, 454-460.

Postman, L. *Psychology in the making*. New York: Alfred A. Knopf, 1962.

Postman, L., & Adams, P. A. Performance variables in the experimental analysis of the law of effect. *American Journal of Psychology*, 1954, *67*, 612-631.

Postman, L., & Adams, P. A. 'Isolation' and the law of effect. *American Journal of Psychology*, 1955, *68*, 96-105.

Postman, L., & Jarrett, R. F. An experimental analysis of 'learning without awareness.' *American Journal of Psychology*, 1952, *65*, 244-255.

Postman, L., & Sassenrath, J. The automatic action of verbal rewards and punishments. *Journal of General Psychology*, 1961, *65*, 109-136.

Razran, G. Semantic and phonetographic generalizations of salivary conditioning to verbal stimuli. *Journal of Experimental Psychology*, 1949, *39*, 642-652. (a)

Razran, G. Some psychological factors in the generalization of salivary conditioning to verbal stimuli. *American Journal of Psychology*, 1949, 62, 247-256. (b)

Rees, H. J., & Israel, H. E. An investigation of the establishment and operation of mental sets. *Psychological Monographs*, 1935, *46*, (6, Whole No. 210).

Russell, W. A., & Storms, L. H. Implicit verbal chaining in paired-associate learning. *Journal of Experimental Psychology*, 1955, *49*, 287-293.

Schroder, H. M. Development and maintenance of the preference value of an object. *Journal of Experimental Psychology*, 1956, *51*, 139-141.

Shepherd, W. O. Learning sets in preschool children. *Journal of Comparative and Physiological Psychology*, 1957, *50*, 15-17.

Shupe, D. R. The assessment of awareness in a verbal conditioning paradigm. *Dissertation Abstracts*, 1966, *26*, 6190.

Sidowski, J. B. Influence of awareness of reinforcement on verbal conditioning. *Journal of Experimental Psychology*, 1954, *48*, 355-360.

Spence, K. W. An experimental test of the continuity and non-continuity theories of discrimination learning. *Journal of Experimental Psychology*, 1945, *35*, 253-266.

Spence, K. W., Homzie, M. J., & Rutledge, E. F. Extinction of the human eyelid CR as a function of the discriminability of the change from acquisition to extinction. *Journal of Experimental Psychology*, 1964, *67*, 343-352.

Spielberger, C. D. The role of awareness in verbal conditioning. *Journal of Personality*, 1962, *30*, (2, Suppl. No. 6), 73-101.

Spielberger, C. D. Theoretical and epistemological issues in verbal conditioning. In S. Rosenberg, (Ed.), *Directions in psycholinguistics*. New York: Macmillan, 1965, pp. 149-200.

Spielberger, C. D., Berger, A., & Howard K. Conditioning of verbal behavior as a function of awareness, need for social approval, and motivation to receive reinforcement. Journal of Abnormal and Social Psychology, 1963, 67, 241-246.

Spielberger, C. D., & DeNike, L. D. Implicit epistemological bias and the problem of awareness in verbal conditioning: A reply to Greenspoon. Psychological Reports, 1963, 12, 103-106.

Spiker, C. C. Experiments with children on the hypothesis of acquired distinctiveness and equivalence of cues. Child Development, 1956, 27, 253-261.

Staats, A. W., & Staats, Carolyn K. Attitudes established by classical conditioning. Journal of Abnormal and Social Psychology, 1958, 57, 37-40.

Staats, A. W., Staats, Carolyn K., & Biggs, D. A. Meaning of verbal stimuli changed by conditioning. American Journal of Psychology, 1958, 71, 429-431.

Staats, A. W., Staats, Carolyn K., & Heard, W. G. Language conditioning of meaning using a semantic generalization paradigm. Journal of Experimental Psychology, 1959, 57, 187-192.

Staats, Carolyn K., and Staats, A. W. Meaning established by classical conditioning. Journal of Experimental Psychology, 1957, 54, 74-80.

Taffel, C. Anxiety and the conditioning of verbal behavior. Journal of Abnormal and Social Psychology, 1955, 51, 496-501.

Tatz, S. J. Symbolic activity in "learning without awareness." American Journal of Psychology, 1960, 73, 239-248.

Thorndike, E. L. The psychology of wants, interests, and attitudes. New York: Appleton-Century, 1935.

Thorndike, E. L. The fundamentals of learning. New York: Teachers College, 1932.

Thorndike, E. L., & Rock, R. T. Learning without awareness of what is being learned or intent to learn it. Journal of Experimental Psychology, 1934, 17, 1-19.

Tolman, E. C. Purposive behavior in animals and men. New York: Century, 1932.

Verplanck, W. S. The control of the content of conversation: Reinforcement of statements of opinion. Journal of Abnormal and Social Psychology, 1955, 51, 668-676.

Verplanck, W. S. The operant conditioning of human motor behavior. Psychological Bulletin, 1956, 53, 70-83.

Verplanck, W. S. Unaware of where's awareness: some verbal operants--notates, moments, and notants. Journal of Personality, 1962, 30, (2, Suppl. No. 6), 130-158.

Wallach, H., & Henle, M. An experimental analysis of the law of effect. Journal of Experimental Psychology, 1941, 28, 340-349.

Wallach, H., & Henle, M. A further study of the function of reward. Journal of Experimental Psychology, 1942, 30, 147-160.

Weinstein, W. K., & Lawson, R. The effect of experimentally induced "awareness" upon performance in free operant verbal conditioning and on subsequent tests of "awareness." Journal of Psychology, 1963, 56, 203-211.

Weiss, R. L., Krasner, L., & Ullman, L. Responsivity of psychiatric patients to verbal conditioning: "Success" and "failure" conditions and pattern of reinforced trials. Psychological Reports, 1963, 12, 423-426.

White, S. H. Age differences in reaction to stimulus variation. In O. J. Harvey, (Ed.), Experience structure and adaptability, New York: Springer Publishing Company, 1966, pp. 95-122.

Wilson, W. C., & Verplanck, W. S. Some observations on the reinforcement of verbal operants. American Journal of Psychology, 1956, 69, 448-451.

Winer, B. J. Statistical principles in experimental design. New York: McGraw-Hill, 1962.

Wong, R., Harrison, J. A., & Stopper, H. Intertrial activity, awareness, and verbal conditioning of children. Psychonomic Science, 1966, 6, 55-56.

Zajonc, R. B. Attitudinal effects of mere exposure. Journal of Personality and Social Psychology Monograph Supplement, 1968, 9, 1-27.

APPENDIX

Dear Parents:

Through the cooperation of the Bellefonte Area School Administration, we have obtained permission to work with a number of children in connection with a program of research being carried out at the University. Specifically, we are interested in the way children learn.

Your child has participated for a few minutes each day for the past week. The study is not designed to tell us anything about a given child, but rather about the age group in general. The procedure resembles a child's game and is very pleasant for the children. A small toy was given to each child for his participation.

The purpose in writing this letter is to inform you about the small present as well as to thank all concerned for permitting us to carry out our study in a cooperative and congenial environment.

Very truly yours,

Richard T. Walls

Richard T. Walls
Department of Educational Psychology
The Pennsylvania State University

Figure 6. Letter of information and thanks to parents, given to S on day 7.

Given: Scores for four independent groups:

1	2	3	4
19	21	40	49
20	61	99	110
60	80	100	151
130	129	149	160

$\underline{p}_{12}$ Computed as follows:

Score 19 is less than each score in Group 2 = 4
Score 20 is less than each score in Group 2 = 4
Score 60 is less than 3 scores in Group 2 = 3
Score 130 is less than no score in Group 2 = 0

$\underline{p}_{12}$ Total 11

m = number of $\underline{S}$s per group = 4

k = number of groups = 4

n = total number of $\underline{S}$s = 16

v = degrees of freedom

$\underline{p}_{12} = 11$

$\underline{p}_{13} = 12$

$\underline{p}_{14} = 13$

$\underline{p}_{23} = 11$

$\underline{p}_{24} = 12$

$\underline{p}_{34} = 12$

$\Sigma pij = 71$

S Computed as follows:

$$S = 2 \sum_{p12}^{p34} - \Sigma m_i m_j$$

$S = 2(71) - (6)\ (4)\ (4) = 46$ (continued)

Figure 7. Sample computation of Jonckheere (1954) S statistic and $\underline{t}$ conversion.

Given: $S = 46 - 1$ if continuity correction is employed.

where $$\mu_2 = \frac{1}{18}\ \{n^2(2n + 3) - \sum_{r=1}^{k} m_r^2\ (2m_r + 3)\}$$

$$\mu_2 = \frac{1}{18}\ \{16^2(32 + 3) - [4(16)\ (2 \times 4 + 3)]\}$$

$$\mu_2 = \frac{8256}{18} = 458.67$$

and $$\gamma_2 = -\frac{36}{25}\ \frac{[k^3(6m^2)+k^2(6m^2+15m)+k(6m^2+15m+10)+6m^2+15m+10]}{(k-1)km[2m(k+1)+3]^2}$$

$$\gamma_2 = -\frac{36}{25}\ \frac{[64(96) + 16(156) + 4(166) + 96 + 60 + 10]}{(3)(16)[8(5) + 3]^2}$$

$$\gamma_2 = -\frac{36}{25}\ \frac{9470}{88752} = -0.15636$$

and $$v = \frac{-3(2 + \gamma_2)}{\gamma_2} = \frac{-3(2 - .15636)}{-.15636} = 36$$

then $$\underline{t}_v = \underline{S}\ \sqrt{\frac{v}{(v+1)\ \mu_2 - \underline{S}^2}}$$

$$\underline{t}_{(36)} = (46 - 1)\ \sqrt{\frac{36}{(37)(458.67) - 45^2}}$$

$$\underline{t}_{(36)} = 2.20$$

Figure 7. Continued.

TABLE XII

NUMBER OF PUPILS PER SCHOOL SECTION (1, 2, 3, 4) OF FIRST GRADE CHILDREN AMONG EXPERIMENTAL CONDITIONS

Experimental Condition	Section			
	1	2	3	4
SS-R4	4	3	2	3
SS-R3	3	3	3	3
SS-R2	2	4	3	3
SS-R1	3	3	3	3
DS-R4	3	3	3	3
DS-R3	2	3	4	3
DS-R2	3	4	2	3
DS-R1	2	3	4	3
SS-EO	4	3	3	2

n = 12

N = 108

TABLE XIII

MEAN AGE IN MONTHS OF Ss IN EXPERIMENTAL CONDITIONS

Stimulus Condition	Rating Condition				All Groups	Control (EO)
	R4	R3	R2	R1		
SS	80.9	81.3	80.9	81.3	81.1	81.8
DS	82.8	81.0	79.5	79.8	80.8	
All Groups	81.9	81.2	80.2	80.6	81.0	

TABLE XIV

ADJECTIVES AND BI-POLAR SCALES USED IN RATING PROCEDURE

Positive Evaluative Adjectives	Negative Evaluative Adjectives	Bi-Polar Adjective Scales
Pretty	Messy	Pretty - Ugly
Best	Ugly	Best - Worst
Good	Worst	Good - Bad
Nice	Awful	Nice - Awful
Happy	Bad	Happy - Mean
Lovely	Mean	Lovely - Rotten
Beautiful	Rotten	Beautiful - Messy

TABLE XV

MEAN FIVE STEP SEMANTIC RATING OF CONCEPTS IN PILOT EVALUATION

Adjective	$\overline{X}$	Adjective	$\overline{X}$
*Messy	4.80	*Nice	1.65
Poor	3.30	*Happy	1.65
*Pretty	1.00	*Awful	4.70
*Ugly	4.90	*Beautiful	1.15
*Best	1.15	Stupid	4.25
*Worst	4.75	Friendly	1.55
*Good	1.15	Healthy	1.25
*Rotten	5.00	Selfish	4.40
*Bad	4.30	*Mean	4.55
*Lovely	1.20	Brave	1.35

N = 20 *S*s

Sample Scale	very good	a little good		a little bad	very bad

---

*Selected for the present study.

TABLE XVI

MEAN DAYS MISSED (INCLUDING WEEKEND) BETWEEN EXPERIMENTAL SESSIONS

Experimental Condition	Session Interim					
	1-2	2-3	3-4	4-5	5-6	6-7
SS-R4	.25	.25	.67	.33	.75	.17
SS-R3	.17	.17	.91	.58	.50	.17
SS-R2	.25	.00	.83	.91	.50	.17
SS-R1	.00	.50	.75	.75	.17	.00
DS-R4	.17	.17	.83	.50	.75	.17
DS-R3	.42	.50	.17	.50	.50	.33
DS-R2	.25	.33	.58	.50	.67	.17
DS-R1	.50	.33	.42	.50	.33	.50
SS-EO	.00	.42	.50	.67	.50	.00

VITA

Richard T. Walls was born in Morgantown, West Virginia on March 29, 1939. He graduated from Morgantown High School in 1957 and attended Bridgewater College in Virginia the following year. He earned a B.S. degree with Honors in Mathematics and Biological Science Education from West Virginia University in 1961 and a M.A. degree from the same institution in Guidance and Counseling in 1963.

Mr. Walls has taught mathematics and engaged in part-time guidance work in the public schools of Morgantown, West Virginia, Miami, Florida, and Cleveland, Ohio for three and one-half years. He has worked as a research assistant with Professor Francis J. Di Vesta from 1964 to 1968 during which time the following research articles have been published:

1. Di Vesta and Walls (1967a)
2. Di Vesta and Walls (1967b)
3. Di Vesta and Walls (1967c)
4. Di Vesta, F. J. and Walls, R. T. Multiple versus single problem training and variations of solution rules in the formation of learning sets. *Journal of Educational Psychology*, 1968, *59*, 191-196.

He is a student member of the American Psychological Association and the American Educational Research Association.

Cognitive Factors in Semantic Conditioning

by

Richard T. Walls

An Abstract of a Thesis

in

Educational Psychology

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

December, 1968

The Pennsylvania State University
The Graduate School
Department of Educational Psychology

ABSTRACT

The experiment was designed to determine the effects of frequency of reinforcement, repeated evaluation of stimuli, and cognitive correlates upon classical semantic conditioning of preference. Previous experimental results found a linear increase in preference for a critical stimulus over four rating days alternated with three conditioning days. An analysis of the earlier work revealed that the effects of conditioning (reinforcements) were inseparable from the possible effects of repeated ratings on the development of preferences.

Accordingly, the present experiment, which ran for seven consecutive school days, was designed to permit the desired analysis of factors affecting the "conditioning" of preferences. The design implied a 2 x 4 factorial analysis of variance. There were two levels of stimulus conditions: at one level the same stimuli were used throughout the experiment; at the other level different stimuli were used on each new conditioning day. The stimulus conditions were crossed with four levels of rating schedules: at one level the Ss rated stimuli on days 1, 3, 5, and 7; those Ss at the next level rated stimuli on days 3, 5, and 7; the third level rated on days 5 and 7; and Ss at the fourth level rated only on day 7. A control group, separate from the main design, rated the same stimuli on all four rating days but received no reinforcements on conditioning days.

A conditioning or training session consisted of a first grade subject spinning a spin-wheel apparatus, on which were printed Greek letters, until 20 reinforcements (marbles) had been received for a

preselected critical stimulus. The rating procedure, on alternate days according to the treatment level, consisted of Ss assigning each of 14 adjectives to any of six stimuli among which was the critical stimulus. The number of negative evaluative adjectives (of seven possible) was subtracted from the number of positive evaluative adjectives (of seven possible) assigned by S to the critical stimulus to provide a preference score.

The replicability, and, accordingly, the reliability of the procedure was supported by reproducing, exactly, the results of the previous study on which the present investigation was based. Preferences were found to increase linearly as a consequence of rating and conditioning when the same stimuli were used. It was suggested that contiguity of the reinforcers and the critical stimulus provided affective contrast with nonrewarded stimuli, thereby increasing the S's attention and effective exposure to the stimulus.

The results also provided support for one of the main hypotheses of the present investigation; preference for the critical stimulus in those groups that rated the same stimuli was a direct function of the number of ratings. It was proposed that this effect was due to increased perception of the elements of the demand characteristics of the experiment during rating sessions and transfer of this knowledge from a rating session to the subsequent conditioning and rating sessions. Although there was some evidence of similar rating effects in "different stimulus" groups, the results were not significant.

The results of awareness measures based on a postexperimental interview complemented the results of a procedure in which S reversed roles with the E. Awareness in the latter procedure was determined by the ability of S to appropriately administer reinforcements to E as he played the spin-wheel game. The role exchange procedure was viewed as a potentially fruitful means for measuring awareness through direct behavioral measures. In general, it was found that Ss who were "conditioned" had both correct reinforcement and correct behavioral hypotheses (awareness of the response-reinforcement contingency).